



ORBIT - Online Repository of Birkbeck Institutional Theses

Enabling Open Access to Birkbecks Research Degree output

Fixing the reference of 'I': immunity to error through misidentification as a guide

<http://bbktheses.da.ulcc.ac.uk/105/>

Version: Full Version

Citation: Orbn, Krisztina (2014) Fixing the reference of 'I': immunity to error through misidentification as a guide. PhD thesis, Birkbeck, University of London.

©2014 The Author(s)

All material available through ORBIT is protected by intellectual property law, including copyright law.

Any use made of the contents should comply with the relevant law.

[Deposit guide](#)
Contact: [email](#)

Fixing the Reference of ‘I’: Immunity to Error Through Misidentification as a Guide

Krisztina Orbán
Birkbeck
University of London

Supervisor:
Professor Dorothy Edgington

A dissertation submitted in
partial satisfaction of the
requirements for the PhD in Philosophy
from Birkbeck, University of London

1 October, 2014

Copyright © Krisztina Orbán, 2014
All rights reserved.

Declaration

I, Krisztina Orbán, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Candidate:

A handwritten signature in dark ink, consisting of a large, stylized 'K' followed by a smaller 'O' and a final flourish.

Krisztina Orbán

Abstract

The first-person pronoun seems to be very special; it has features which other kinds of singular terms lack. Uses of 'I' have guaranteed right reference (GRR), they cannot lack referent and a user of 'I' cannot use 'I' for the wrong object, and some uses of 'I' have immunity to error through misidentification (IEM). For an immune self-ascription made on a basis (like proprioception) the subject cannot misidentify who is *F*. I will elaborate in great detail how to characterise these essential features: IEM and GRR. Peacocke and Campbell suggested a Reference-fixing test for accounts of 'I'. According to this test what fixes the reference of 'I' has to be able to account for the essential features of 'I' (GRR and IEM). I will use this test to evaluate the view I will propose, which I call the Simple View.

I suggest that the subject uses 'I' for the object which she knows through internal ways of knowing. Examples of internal way of knowing are proprioception, introspection and nociception (pain perception). What fixes the referent of 'I' is simply that the subject use 'I' for the object which she knows from the inside through internal ways of knowing. This explains how GRR is possible: why the subject cannot be wrong about the referent of 'I'. And this explains IEM, why based on certain ways of knowing (internal ways of knowing) the self-ascription is such that who has the property cannot be misidentified. I try to show that other ways to fix the referent cannot provide the most fundamental explanation of 'I' because none of them can account for the essential features of 'I'.

Contents

List of Figures and Tables	6
Guide to Notation and List of Abbreviations	7
Acknowledgements	8
 Part I	
Introduction	10
1. Wittgenstein on Immunity:	20
How does the Subject Know of the Referent of 'I'?	
2. Anscombe's Challenge: Guaranteed Right Reference	31
3. The Basic Characterisation of IEM	46
Appendix: Departing from Pryor	56
4. The Myth of Absolute Immunity	59
5. Is immunity to error through misidentification	77
immune to empirical counterexamples?	
6. On the Identification-freedom Explanation of IEM	92
Part I: The Two Kinds of IEM: Evans's Identification-freedoms	93
Part II: Varieties of identification: Two kinds of identification	106
Appendix: Identification-dependent Immunity?	124
7. The General Explanation of IEM and	128
the Reference-fixing Relation	
 Part II	142
8. The Simple View: Fixing the Reference of 'I'	143
9. The Simple View: The Test	165
10. Objections to the Simple View	184
 Conclusion	216
Bibliography	219

List of Figures and Tables

Figure 1. Self-identification and Frege-identification	109
Figure 2. Varieties of identifications: the examples	115
Figure 3. Object-identification and Frege-identification	116
Table 1. Constraints	163
Table 2. Features of the first-person pronoun	166
Table 3. List of Objections	184

Guide to Notation and List of Abbreviations

IEM – immunity to error through misidentification

Immune – immune to error through misidentification

GRR – guaranteed right reference

GRR1 – guarantee against lack of reference

GRR2 – guarantee against incorrect reference

Fa (B) – a is F based on a basis B , where ' F ' is a predicate, ' a ' is a singular referring term and ' B ' is a particular basis

Fi – I am F

B – basis: a way of knowing like proprioception, introspection, vision, or testimony

W – a way of knowing, like proprioception, introspection, vision, or testimony

OI – Object-identification

FI – Frege-identification

IF – Identification-freedom

Acknowledgments

I am enormously grateful to my supervisor, Dorothy Edgington, whose vigilance and sharp mind helped me through. Her tutelage straightened my style and my thinking. She not only supervised me, but also provided personal support and warmth. The trenchant criticism of Hong Yu Wong was essential for my work in many respects. His theory of action and bodily awareness inspired my work and directed my attention to what it is to know of my body from the inside. I had very helpful discussions with the psychologists Denis Proffitt and Wolfgang Prinz about the Simple View. I wish to thank Francois Recanati, John Campbell, Caleb Liang, and Timothy Lane for discussions about some of the other chapters. I am grateful to the audience in my talks in Lausanne, Budapest, Essen, Taipei, Santiago de Compostela, London, Porto, Granada, Tübingen, and Noto for questions and comments. I was very lucky to get very good feedback on many occasions where I presented parts of my half-baked works for the PONS group in Tübingen. I thank the members of this group: Chiara Brozzo, Gregor Hochstetter, Ferdinand Pöhlmann, Eva Engels, Alex Morgan, Katia Samoilova, and Hong Yu Wong. I also wish to thank Julia Peters, Victor Verdejo, Victor Gan, Cristina Borgoni, Christoph Michel, Elvira di Bona, Paul Snowdon, Thor Grünbaum, Daniel Morgan, Manuel Garcia-Carpintero, Jen Hornsby and Doug Lavin for discussion. Finally, I wish to thank my parents and my better half. They provided warmth and happiness. I especially like to be awakened by the sound of the violin.

Part I

Ambrose Bierce has a pleasant entry under "I" in the Devil's Dictionary:

In grammar it is a pronoun of the first person and singular number. Its plural is said to be We, but how there can be more than one myself is doubtless clearer to the grammarians than it is to the author of this incomparable dictionary. (Anscombe 1975).

Each of us can fix the referent of the word 'I' by means of acquaintance with oneself, self-acquaintance. (Kripke 2011a: 301).

Introduction

The aim of this thesis is to understand how the reference of 'I' is fixed. I will propose how and to which object the reference of 'I' is fixed without assuming what essential kind of object it is. In particular, I propose that the subject knows through internal ways of knowing (like proprioception and introspection) of the object which she is, and uses 'I' for this object. As I will argue, the subject is directly acquainted with the object which she is through internal ways of knowing. Only a unique object, the subject, is present to the subject through internal ways of knowing. *The subject uses 'I' for this unique object; this fixes the referent of 'I'.* My proposal will be called the **Simple View**. My main concern is to answer the question 'How is the referent of 'I' fixed?' My contention is that the Simple View provides the most fundamental answer.

1. The strategy

In the first part of the thesis (chapters 1-7), my concern is elaborating what the essential features of the use of 'I' are (including guaranteed right reference and immunity to error through misidentification). In the second part of the thesis (chapters 8-10), I will try to show that only the Simple View can explain the essential features of 'I'. Consequently, the Simple View provides the most fundamental answer to the question 'How is the referent of 'I' fixed?'

It seems that the first-person pronoun is special. It has features which other singular terms and even other essential indexicals ('now', 'here') lack. For example, uses of 'now' and 'here' lack a guarantee against incorrect reference. I can think that my current location in time or space is different from what it actually is. Thus I can be incorrect about what the referent of my use of 'here' or 'now' is. However, I cannot try to refer to myself but fail to refer to myself by using 'I' or refer to the incorrect thing. It is only plausible to think that what fixes the referent of 'I' should explain those features which make the first-

person pronoun different from other referring terms (Peacocke 2008: 92, 2012: 144, Campbell 2012).

When one's project is to understand how the reference of 'I' is fixed, then one has to consider the essential features of the first-person pronoun. I take the essential features of the use of 'I' to be constitutive for understanding the first-person pronoun.

What are the essential features of uses of 'I'? Building on the work of Peacocke (2008: 92 and 2012) and Campbell (2012) we can find two strong candidates for essential features of 'I': guaranteed right reference (hence **GRR**) and immunity to error through misidentification (hence **IEM**). I discuss why each is an essential feature in turn.

Guaranteed right reference is agreed to be essential for understanding the first-person pronoun.¹ Uses of 'I' have a double guarantee, **guaranteed right reference**: (i) they cannot lack reference and (ii) they have a guarantee against incorrect reference (hence **GRR2**). All uses of the first-person pronoun are supposed to have guaranteed right reference and only the first-person pronoun has this feature.²

It seems to be hard to answer how a guarantee against incorrect reference is possible. But this is an easy question for the Simple View. Nothing but the subject can be known through internal ways of knowing. Accordingly, the subject cannot be wrong about which object she knows through such ways of knowing, and so which object she uses 'I' for. This enables the subject to know of the object which she is without the possibility of mistake as the guarantee against incorrect reference requires.

IEM is a feature of some self-ascriptions made on a certain basis – as I will argue. A self-ascription, I am in pain, based on feeling pain, is **immune to error through misidentification** (hence, for the adjectival form, *immune*) because, on this basis, feeling pain, the subject cannot misidentify who is in pain. IEM is only a feature of certain but not all self-ascriptions. I will argue in chapter 7 that

¹ There are some dissenters. Some deny GRR because they deny that 'I' refers (e.g. Wittgenstein and Anscombe), and some deny it for cases when one is a brain in a vat (e.g. Evans). *Pace*

² For 'now' and 'here', the subject can intend to refer to the wrong location in time or space. I can intend to refer to Paris by using 'here' when I am in Budapest and I can intend to refer to tomorrow using 'now' when I am mistaken which day it is today.

the general explanation of IEM involves the reference-fixing rule. In chapter 8, I will argue that the reference-fixing rule explains why and when a self-ascription is immune. I will show that because the reference of 'I' and demonstratives are fixed differently there is a difference between first-person IEM and demonstrative IEM (chapter 6).

It is easy to propose a view. What is hard is deciding whether it is correct. It would be best if we could find a solid test that a view has to pass in order to be accepted as the fundamental explanation of how the reference of 'I' is fixed. If the Simple View were the only view that could pass the test, then it would appear to be the most fundamental explanation of the phenomenon (from amongst the examined views). In the literature, we can find such a test for the reference-fixing rule of the first-person pronoun, suggested independently by Peacocke (2008: 92) and Campbell (2012). The idea is straightforward. If we wish to have a theory of how the reference of the first-person pronoun is fixed, then this theory should account for the essential features of the uses of the 'I', especially GRR and IEM. Call this the **Reference-fixing Test**. I will use this to test the Simple View in chapter 9.

For this project to work, we will have to have a precise understanding of GRR and IEM. Consequently, the thesis is written in two parts. In the first part of the thesis, I will concentrate on getting a firm handle on IEM and GRR. In the second part of the thesis, I will develop the Simple View and test whether it can explain IEM and GRR. I will argue that only the Simple View can pass the Reference-fixing test.

2. Conceptual assumptions and preliminaries

I find it plausible that the subject refers to a particular object by using 'I' and all uses of 'I' have GRR and some judgments are immune. I have not found any sound challenge to 'I' being a referring term so I will simply accept that 'I' refers.

It is plausible that the subject, whatever essential kind it falls under, should be embodied (Shoemaker 1984c, Wiggins 2001, Burge 2010, Snowdon 2014). Thus, 'I' refers to a physical object. I treat the main question of the thesis ('How is the referent of 'I' fixed?') as distinct from the main question of personal

identity ('Which essential kind are we?'). I try to be as neutral on the second question as possible.

One of the leading themes of the thesis is why and how the basis of a judgment matters for whether it is immune. When I judge something it could be based on different bases like vision or touch. 'Basis' means a ways of knowing through which the subject acquires knowledge. Accordingly, the basis of a judgment could be vision, touch, audition, olfaction, proprioception, kinaesthesia, introspection, inference or testimony. A judgment is not IEM *per se*, but immune relative to a certain basis; it is not immune relative to all possible bases. I call this **basis-relativity** (argued for in chapters 1-4). Basis-relativity is an important point for me because I try to show in chapters 7-8 that a judgment about *a* is immune relative to '*a*' on a basis B iff *a*'s reference is fixed through this very basis B.

Let me introduce some notation that will be used throughout the thesis to express judgements and their basis-relativity. I will abbreviate the claim '*a* is *F*' as ***Fa***. In '*Fa*', '*F*' is a predicate and '*a*' is a singular referring term, it could represent the referent of '*I*'. '*I am F*' will be abbreviated as ***Fi***. When I want to express that a judgement *Fa* is made on a certain basis B, I will abbreviate '*Fa*, relativized to a certain basis B' as ***Fa (B)***.

What is the item that can be qualified as being immune to error through misidentification? It seems that there is no agreement on this front in the literature. The only thing which is agreed by all parties is that something like *Fa* relative to '*a*' would be immune. Variations are plenty. The majority of authors relativise the judgment to its basis, while a minority of authors think that a judgment independent of its basis can be immune.³ In my discussion, I will use 'thought' and 'judgment' interchangeably as candidates for the item that is immune on a certain basis. I try to be as neutral as possible about specific

³ Just to take some examples. For Wittgenstein, a *self-ascription* is immune *per se* (without relativizing it to its basis). For Shoemaker, Campbell and O'Brien a *basis-relative judgment/self-ascription*, *Fa (B)*, is immune. For Evans, *knowledge of a basis-relative proposition/judgment*, *Fa (B)*, is immune. According to Peacocke, a *basis-relative thought relative to a Circumstance*: *Fa (B, C)*, is immune. Higginbotham holds that a thought, *Fa*, is immune. For Recanati, a *basis-relative thought/judgment/belief* is immune, *Fa (B)*. I assume that these differences depend on the framework and on what else the theorist assumes.

technical and ontological details and to be fairly independent from any particular framework.

Below I will sketch each chapter of the thesis.

3. The chapters

Chapter 1. Wittgenstein on Immunity: How does the Subject Know of the Referent of 'I'?

In chapter 1, I will discuss Wittgenstein's original characterisation of IEM. Wittgenstein distinguishes between two kinds of self-ascription: (i) 'I' used as subject and (ii) 'I' used as object.⁴ The two uses exclude each other and are introduced through examples. Wittgenstein provides a criterion to distinguish the two sets of examples: whether self-ascriptions are recognition-based. 'I' used as object is recognition-based, while 'I' used as subject is not.

Wittgenstein assumes that recognition is necessary for the subject to know of the physical object which she is; recognition enables the subject to refer to a physical object. Subject uses of 'I' are recognition-free. Thus subject uses of 'I' do not refer. I will show that the subject is able to know of the physical object which she is without recognition (in the sense that she need not single the object out from among many).

Against Wittgenstein, I will argue that recognition-freedom is not sufficient for IEM. Additionally, a self-ascription is not immune *simpliciter*, but only immune relativized to its basis. Throughout the first half of the thesis I will argue for the claim that IEM is basis relative which Wittgenstein denies.

Chapter 2. Anscombe's Challenge: Guaranteed Right Reference

In chapter 2, I will discuss GRR as presented in Anscombe's paper 'The first person' (1975). However, in elaborating a more accurate understanding of GRR, I will move away from the original formulation by Anscombe.

I will explicate the guarantee against incorrect reference (GRR2) by employing a distinction introduced by Kripke (1977). On this picture, the

⁴ Wittgenstein's notion of 'use of 'I' as subject' is different from Kant's notion of 'consciousness of oneself' as subject. Wittgenstein only discusses two different kinds of self-ascriptions while Kant discusses transcendental presuppositions of thinking in general. I return to the Kantian notion briefly in chapter 5 and in greater detail in chapter 10.

subject, by using 'I', cannot intend to refer to the wrong thing because the speaker's referent and the semantic referent cannot come apart. This new notion will be called **aboutness-error freedom** and will provide the explanation of the guarantee against incorrect reference.

For Anscombe, GRR is essential for *referential* uses of 'I'. But, as she argues, there is a circumstance when, for a token 'I', GRR cannot be satisfied. GRR excludes the possibility of mistakenly taking the wrong object to be me. This is not possible unless the subject is directly acquainted with the object, as Anscombe argues. (The relevant object has to be physically present to the subject.) Anscombe uses the sensory deprivation tank (where all the subject's senses are suspended) to argue that 'I' cannot be a referring term (used referentially) because the subject might use 'I' without being acquainted with its referent.

Anscombe provides a challenge to the Simple View. The Simple View requires that whenever the subject can use 'I' she is directly acquainted with the object, herself. Anscombe's challenge is how direct acquaintance is possible in the sensory deprivation tank when the subject uses 'I'. Call this **Anscombe's challenge**. I will return to answer it in the last chapter.

Chapter 3. The Basic Characterisation of IEM

In chapter 3, I will characterise the basic notion of IEM through explicating the notion of misidentification that is involved in IEM. As I will argue, a judgment, *Fa*, based on basis *B*, is immune iff the subject cannot misidentify, on this basis, *B*, that, this object, *a*, is *F*. Accordingly, misidentification happens when the subject thinks *Fa* on a basis, *B*, which is for thinking *Fb*. Pryor (1999) offers an alternative to this standard notion of misidentification: which object misidentification. I will argue that the standard notion of misidentification does a better job at characterising IEM.

The basic notion of IEM requires basis-relativity; the judgment has to be relativized to its basis. An immune judgement cannot be misidentified on a particular basis. This will be argued for conclusively only in the next chapter.

Chapter 4. The Myth of Absolute Immunity

In chapter 4, I will discuss Shoemaker and his introduction of IEM. We owe the distinction between absolutely immune as opposed to circumstantially immune judgments to him. A judgment is **absolutely immune** if it is immune in all possible circumstances (where it could be uttered) while a **circumstantially immune** judgment is only immune in some but not in all circumstances. Accordingly, if *Fa* is absolutely immune then it is immune on all bases. Basis-relativity (as I construe it) entails that the thought (B) is immune on B but it is *not* immune on all bases. Shoemaker denies that absolute IEM is basis-relative.

I will argue that a judgment is not IEM *per se* (basis-independently), but immune relative to a certain basis. It is immune relative to certain but not all bases. This idea – basis-relativity – can be found in Evans, but he only uses an example to demonstrate the point. I am obliged to argue for it, for the reason that this point is crucial for working out the general explanation of IEM (chapter 7) and the Simple View, where the reference of ‘I’ is fixed through some distinctive bases: internal ways of knowing (chapters 8-10).

IEM could only be a basic explanandum if it is not a mere consequence of another feature, like infallibility. A judgment is **infallible** if whenever it is uttered it is true. If a judgment ‘I am *F*’ is infallible it follows from its infallibility that it is immune. This would not be a paradigmatic case of IEM but a paradigmatic case of infallibility.⁵ We have to focus on the correct explanatory target: IEM.

To get the explanatory target right, I will work out the **self-sufficiency principle**. According to this principle, IEM cannot simply follow from another feature like infallibility or self-intimacy. A state, like being in pain, is **self-intimating** if whenever the subject is in pain, she knows that she is in pain. I will argue that if we accept the self-sufficiency principle then we have to accept basis-relativity. If we deny the self-sufficiency principle then we have to change the explanatory target from IEM to something else, like infallibility, from which basis-independence follows.

⁵ “I am thinking that *p*” based on introspection is infallible. For this reason, it cannot be a paradigmatic case of immune self-ascription.

Chapter 5. Is immunity to error through misidentification immune to empirical counterexamples?

In chapter 5, I will defuse some empirical cases which have been claimed to be putative counterexamples against the phenomenon of the IEM of self-ascriptions. I will focus on three cases: thought insertion (Campbell 1999b), somatoparaphrenia, and the body swap illusion (both discussed in Lane and Liang 2011).

Chapter 6. On the Identification-freedom Explanation of IEM

In chapter 6, I examine the standard understanding of IEM due to Evans. Evans proposes to explain IEM in terms of Identification-freedom. I will show that Evans's picture is far from being as simple and uniform as it is supposed to be. I will make an important distinction between two kinds of Identification-freedom found in Evans's work: Factual and Strong Identification-freedom (Evans 1982). Factual Identification-freedom requires only that the thought, *Fa*, is not *in fact* grounded on an identification ($a=b$), while Strong Identification-freedom requires that on the basis, *B*, on which the thought, *Fa*, is based, *Fa* **cannot be** grounded on an identification ($a=b$).

Why would Evans work out two different explanations, factual and strong identification-freedom, for immunity? I argue that the immunity of self-ascriptions cannot be explained in the same way as the immunity of perceptual demonstratives. In Evans, we find that one explanation is for explaining the IEM of self-ascriptions (strong identification-freedom) and another explanation is designed for explaining the IEM of demonstrative thoughts (factual identification-freedom). I conclude that the difference in the explanation occurs because there are two different kinds of IEM (first-person IEM and demonstrative IEM).

In the second part of this chapter, I will show that the notion of identification is not univocal. I will differentiate two different notions of 'identification' and argue that what opens the possibility of misidentification for perceptual demonstrative thought does not open this for self-ascription and vice versa.

Chapter 7. The General Explanation of IEM and the Reference-fixing Relation

It seems that the IEM of self-ascriptions and perceptual demonstrative thoughts are in some sense different and in some sense similar. In chapter 7, I will look into what is shared between the two kinds of IEM. The general explanation of IEM – as I will argue – will rely on the reference-fixing relation. According to the general explanation of IEM a judgement *Fa* (B) is immune iff the referent of ‘a’ is fixed through B. The reason why we have two kinds of IEM derives from differences on how the referent of the first person pronoun and a perceptual demonstrative is fixed.

Chapter 8. The Simple View: Fixing the Reference of ‘I’

In chapter 8, I will investigate how the reference of ‘I’ is fixed. As I will argue, the subject knows of the object which she is through distinctive ways of knowing which are such that only the subject can be known through such ways of knowing. Proprioception and introspection are the most obvious examples of such special ways of knowing through which only the subject and her states, events and processes can be known. Through such special ways of knowing, from the inside, the subject cannot be mistaken about which object she knows of.

The subject simply uses ‘I’ for the object of which she knows from the inside. The subject cannot be mistaken about which object this is because she is directly acquainted with it and it is the only object which she can know from the inside.

GRR is an essential feature of uses of ‘I’. We try to understand how GRR is possible. What are the necessary constraints for making a guarantee against incorrect reference possible? My task is to find some special ways of knowing which allow the subject to know of the object which she is in a way that excludes the subject from using ‘I’ for the wrong object. I discuss which features of proprioception and introspection make the guarantee against incorrect reference possible.

According to my proposal, the Simple View, the referent of ‘I’ is fixed by the subject’s use of ‘I’ for this object which she is directly acquainted with and of

which she cannot be mistaken about which one it is. This is the object with which the subject is acquainted with through special ways of knowing – from the inside. For the subject, four constraints will be specified as jointly necessary and sufficient for providing direct acquaintance with the object which is the subject. The guarantee against incorrect reference is possible because of the distinctive features of the ways of knowing through which the referent of ‘I’ is fixed.

Chapter 9. The Simple View: The Test

In chapter 9, I will examine the essential features of uses of ‘I’. We have learnt from Peacocke and Campbell that the test for the Simple View is whether it can explain the essential features of the uses of ‘I’ (GRR and IEM). I use GRR2 as a guide in isolating the constraints on what makes the ways of knowing special (chapter 8). Thus, unsurprisingly, the Simple View explains GRR2 very well. My strategy will be to show that unless the referent of ‘I’ is fixed as the Simple View suggests, it is not clear what can account for first person IEM and GRR.

Chapter 10. Objections to the Simple View

I will consider several objections to the Simple View and answer them. I hope at the end of the thesis the Simple View will not only provide a suitable explanation of why all uses of ‘I’ have GRR and why some uses are immune, but it will also show why the first-person pronoun is special and different from any other singular referring term.

Chapter 1

Wittgenstein on Immunity:

How does the Subject Know of the Referent of 'I'?

Abstract

Wittgenstein provided the first characterisation of the IEM of self-ascriptions. For Wittgenstein, recognition-freedom is necessary and sufficient for IEM of a self-ascription. Furthermore, he holds that a self-ascription is immune *simpliciter*. Against Wittgenstein, I will argue (i) that recognition-freedom is not sufficient for IEM and (ii) that a self-ascription is not immune *simpliciter*. I will use Evans's criticism of Wittgenstein as a starting point to argue for the claim that IEM is basis-relative. (My argument will continue in chapters 3-4.)

To understand why IEM can teach us something about the first-person pronoun, I have to answer the following question: Why would recognition tell us anything about the referentiality of 'I'? Wittgenstein, it appears, was puzzled about how the subject knows of the physical object which is herself without recognising that object. We will find that the subject has a special way of knowing of herself which does not require recognition (in the sense that I need not single the object, the referent of 'I', out from among many).

1. Preliminaries

Immunity to error through misidentification is always related to the understanding of the first person pronoun and self-knowledge. I will try to show why IEM is indispensable for understanding the first-person pronoun. The best place to start our discussion is where the debate starts historically: Wittgenstein's treatment of IEM. Wittgenstein is only interested in immunity because it apparently provides support for thinking that we are Cartesian Egos

– he wants to deny this. What motivates Wittgenstein’s discussion is his desire to show that we have no reason to think that we are Cartesian Egos. His surprising solution is this: in immune self-ascriptions, where ‘I’ used as a subject, ‘I’ does not refer. After two pages describing the phenomenon of immunity, such a radical conclusion does not sound very plausible.

I am not interested in arguing for the referentiality of the first person pronoun against Wittgenstein, who wished to deny it. *Pace* Wittgenstein I simply assume that the subject by using ‘I’, without exception, refers and intends to refer herself.⁶ I assume that all uses of ‘I’ refer to a *material object* with both mental and physical qualities as opposed to a Cartesian Ego (Strawson 1959, Shoemaker 1968, Burge 2010, Peacocke 2014, Snowdon 2014). This provides us with an exciting question, which is how IEM is possible, if ‘I’ refers. Is there something about how the subject knows of the referent of ‘I’ which makes self-ascriptions less vulnerable to certain kind of mistakes – such as misidentification?

2. Immunity to error through misidentification

The label **immunity to error through misidentification** was introduced by Shoemaker in ‘Self-Reference and Self-Awareness’, but the phenomenon of IEM was known before. Though the phenomenon only came to general philosophical attention with Shoemaker’s article, the discovery of the phenomenon is credited to Wittgenstein, because of his discussion in the *Blue and Brown Books*.

The paradigmatic example of IEM is the self-attribution of pain. For ‘I am in pain’, based on feeling, the subject cannot be wrong about *who is in pain*. In contrast, based on vision, looking at the mirror, the subject can misidentify *who has messy hair* when she judges that ‘I have messy hair’.

⁶ By ‘use of ‘I’’, I will always mean a literal use of ‘I’ where the speaker/thinker intends to refer to herself. For example, in the following sentence the first occurrence of ‘I’ is a non-literal use while the second occurrence is a literal use of ‘I’. If I were you then I would be where I am not. This could not make any sense if there were not two different uses of ‘I’ in it; but it makes sense. A subject cannot use ‘I’ unless she knowingly and intentionally refers to herself (Anscombe 1963, Castaneda 1966, Evans 1982, O’Brien 2007, Peacocke 2008, 2014).

3. Wittgenstein's two theses concerning IEM

3.1. The first thesis

To introduce the phenomenon Wittgenstein presents us with two separate sets of self-ascriptions. The first set contains the paradigmatic cases of 'I' used as subject, while the second set provides the paradigmatic cases of 'I' used as object.

Now the idea that the real I lives in my body is connected with the peculiar grammar of the word "I", and the misunderstandings this grammar is liable to give rise to. There are two different cases in the use of the word "I" (or "my") which I might call "the use as object" and "the use as subject". Examples of the first kind of use are these: "My arm is broken", "I have grown six inches", "I have a bump on my forehead", "The wind blows my hair about". Examples of the second kind are: "I see so-and-so", "I hear so-and-so", "I try to lift my arm", "I think it will rain", "I have [a] toothache". (Wittgenstein 1958: 66-67)

I will distinguish between two theses Wittgenstein advanced. The first thesis demarcates the two uses of 'I'. The object use of 'I', according to Wittgenstein, is different from the subject use of 'I' as it involves recognition of a particular person by his bodily characteristics.⁷

One can point to the difference between those two categories by saying: The cases of the first category [the object use of 'I'] involve the recognition of a particular person (...). (Wittgenstein 1958: 66-67)

We feel then in the case in which "I" is used as a subject, we don't recognize a particular person by his bodily characteristics (...). (Wittgenstein 1958: 69).

⁷ Wittgenstein's notion of 'use of 'I' as subject' is different from Kant's notion of 'consciousness of oneself' as subject. In the Kantian tradition, the self can be presented as an object or as a subject. Wittgenstein only demarcates two sets of self-ascriptions from each other. In contrast, the Kantian notion (e.g. Cassam 1997) applies not only to self-ascriptions, but to all thoughts of the subject. ('I think...' can be attached to all thoughts of the subject.) In chapter 10 I will return to this question.

We can formulate Wittgenstein's idea as follows:

WT1. The object use of 'I' involves recognition of a particular person, while the subject use of 'I' does not involve recognition of this kind.

We turn to the second thesis.

3.2. The second thesis: subject uses of 'I' are immune

According to Wittgenstein, subject uses of 'I' are immune and object uses of 'I' are not immune. The reason for this is that subject uses of 'I' are recognition-free, while object uses of 'I' are recognition dependent. 'Recognition' is always used for singling the object, the referent of 'I', out from among many objects in some way (e.g. by bodily characteristics).

One can point to the difference between those two categories by saying: The cases of the first category [i.e. the object use of 'I'] involve the recognition of a particular person, and there is in these cases the possibility of an error, or as I should rather put it: The possibility of an error has been provided for. (...) On the other hand, there is no question of recognizing a person when I say I have [a] toothache. (Wittgenstein 1958: 66-67).

Once again, we can formulate Wittgenstein's idea as a thesis:

WT2. A self-ascription is immune iff it only involves subject uses of 'I'.

For Wittgenstein, just and only recognition of an object opens the possibility of misidentification. Recognition-freedom is necessary and sufficient for IEM. Wittgenstein holds that the subject use of 'I' does not refer because it is recognition-free. This presupposes that one cannot know of a present physical object, if that object was encountered before, unless one recognises that object. Against Wittgenstein, I will argue that recognition-freedom is not sufficient for

IEM and that the subject can know of a present physical object without recognising it. I start with the first claim.

4. Recognition-freedom is not sufficient for IEM

4.1. The Mistaken Identity Example

Consider the following example. Suppose that in 1949 I have amnesia and have forgotten who I am. But I am deceived by my friends and made to believe that *I am Anscombe*. I know that Anscombe was a student of Wittgenstein and I am sure of this. But I know that like me my friends like to play tricks – even so this happens only occasionally, and I usually trust them. In this case my friends have managed to convince me that I am Anscombe, and on the basis of the presuppositions above, I judge based on testimony and inference:

J1) I was Wittgenstein's student.

I might meet with Wittgenstein who insists that he has never seen me before. He looks surprised by my suggestion that I was his student. I ask whether Anscombe was his student. He agrees with a cheerful, if not an ironic smile. Consequently, I conclude that I cannot be his former student; I cannot be Anscombe. There is no recognition involved. I do not recognise myself as Anscombe, I only believe that I am Anscombe. Still I can misidentify who was Wittgenstein's student, so J1 (B=testimony, inference) is not immune. We can see that misidentification is involved because if I doubt that I am Anscombe and that I am Wittgenstein's student, then I should believe that Anscombe is his student and not me.

Here is another case. During the socialist period of Hungary, I was in the underground. I seem to remember that I came up with the idea of holding a peace demonstration on the theme of "The funeral of the known soldier" (as opposed to the usual commemoration of unknown soldiers). There were several other small group demonstrations at this event. All the evidence points to the fact that it was my friend who came up with this scheme to put a straw man in military uniform and I did not add any ideas to the scheme. It seems I misidentified who came up with that splendid scheme which was the most

successful of all small group demonstrations at that time. But recognition does not seem to be involved. I still seem to remember that I came up with the idea – but I am wrong about who came up with the idea.

I have only tried to give a flavour of how misidentification of *who is F* seems to be possible without recognition. For Wittgenstein, just and only recognition of an object should open the possibility of misidentification. As we have seen, this cannot be correct.

4.2. The head-knocking example

Let me examine whether Wittgenstein's examples of subject and object uses of 'I' provide us with a better understanding of what could be considered as immune.

Let us look at Wittgenstein's example of an object use of 'I'.

J2) I have grown six inches.

According to Wittgenstein, J2 is **not** immune. But J2 can be immune even on an observational basis, as we will see. Suppose I knock my head on a doorframe which I know is six inches taller than me. If I assert J2 just and only because I knocked my head to the believed-to-be-six-inches-taller-than-me doorframe, then on this ground there is no room to think that if I did not grow six inches then someone else did. Thus, contrary to Wittgenstein, on this particular ground, J2 is immune relative to 'I'. According to Wittgenstein, J2 requires recognition and recognition opens the possibility of misidentification. But in our example J2, so judged in the situation, does not and cannot involve recognition in its ground.

One might think that the door might shrink six inches. This might happen, for example, as a result of an earthquake. But still nothing but I can grow six inches – if anyone has grown. We have to specify the basis in order to evaluate whether the judgement is immune.

5. Basis matters – the leading theme

There are in fact two different ways that Wittgenstein differentiated the two uses of 'I' from each other. The first introduction was through listing paradigmatic examples and the second introduction was by answering whether the judgements are based on recognition. Neither of them allows that a judgment might be immune on one but not immune on another basis. This does not seem to be correct, as Evans observed:

Wittgenstein's discussion does not take sufficient account of the fact that the property of immunity to error through misidentification is not one which applies to propositions simpliciter, but one which applies only to judgments made upon this or that basis. (Evans 1982: 218-219).

Evans pointed out what is lacking in Wittgenstein's account of IEM. Evans argues that a judgment like 'I have my legs crossed' is only immune relativized to its basis. Suppose I judge based only on proprioception:

J3) I have my legs crossed.

Proprioception is a perceptual modality that provides me with a sense of the position of my body parts relative to one another from the inside. When it turns out that my legs are not crossed, then, on this basis, I cannot judge that someone else's legs are crossed. On this basis, I cannot misidentify who has her legs crossed, thus J3, based on proprioception, is immune. However, when I judge J3 based on vision, then the case is different. If it turns out that my legs are not crossed then I can judge on the same basis that 'this person's (whom I see in the mirror) legs are crossed' (Evans 1982: 218-224). We might worry that if J3 is based both on proprioception and vision, then the one who has her legs crossed can be misidentified. Thus we suggest that J3 solely based on proprioception is immune. This can be repeated for other judgements like

J4) My arm is broken.

or

J5) I have a bump on my forehead.

If I judge J4 based solely on feeling pain without exploiting vision, then I cannot misidentify whose arm is broken. Similarly, if I feel a stretching feeling from the inside – what I feel when my bump due to knocking my head is swelling – I might judge, based only on this, that I have a bump on my forehead. If my basis for J5 is only what I know from the inside then I cannot go wrong about who has a bump on her forehead.

One might worry that this is only restricted to some examples so it might not be legitimate to generalise from them. This is correct, so we only conclude that at least some self-ascriptions are such that the self-ascription made on a certain basis is immune. I will argue in the following chapters for the generalisation of this claim. It seems that contrary to what Wittgenstein claims, the judgment has to be relativized to its basis for deciding whether it is immune. For IEM, basis seems to matter.

The subject use of 'I' is supposed to mark out the class of self-ascriptions which are immune. Recognition-freedom does not seem to delineate this class very well. But what is the big picture behind Wittgenstein's investigation? How does Wittgenstein arrive at the conclusion that subject uses of 'I' do not refer from their lack of recognition of an object?

IEM is always considered essential for understanding the first-person pronoun. But it has never been made clear why IEM matters for understanding the first-person pronoun. Let me elaborate on the background of Wittgenstein's argument so as to see what IEM can teach us about the first-person pronoun.

6. How does the subject know the referent of 'I'?

Wittgenstein's second thesis characterises IEM: a self-ascription is immune iff it only involves a recognition-free subject use of 'I'. But what is recognition? On a plausible reading of Wittgenstein, recognition of somebody (a physical object) happens by identifying an object by its bodily characteristics (Wittgenstein 1958: 69). Ordinary instances of identifying an object by its bodily

characteristics would include cases of recognising someone by some relatively permanent bodily characteristics, like a birthmark, an unusual smile, always having a strange posture, or something more holistic, such as having a 'Johnny-like look'. What enables me to recognise someone is the comparison between how she usually looks and how a person looks now.

However we have to realise that it is not these relatively permanent bodily characteristics which enable the subject to recognise herself. I want to note that even though we see others and we might recognise them by their bodily characteristics, we do not see our face except with the aid of mirrors or reflecting surfaces like lakes. We know our bodily characteristics by looking at mirrors. Accordingly, I cannot identify myself by my typical bodily characteristics – e.g. how my face or I would look – unless I learn them through looking at mirrors. But I first have to know that it is I who I see in the mirror before I can learn what my bodily characteristics are. So it seems that the recognition approach cannot provide us with the correct direction for uncovering how the subject knows of the referent of 'I'.

6.1. The big picture:

The non-referentiality argument and its presuppositions

There is a big picture behind Wittgenstein's discussion about IEM concerning whether 'I' refers and, if it refers, how it can refer to a physical object. Wittgenstein uses IEM to argue that, in some uses of 'I' (the subject uses), 'I' does not refer. Let me call this the **non-referentiality argument**.

The argument is as follows. For Wittgenstein, if 'I' refers, it refers to the physical object which always has to be present to the subject, and the subject has to know of that object when she uses 'I'. Let us accept this for the sake of argument.⁸ However Wittgenstein claims that without recognition one cannot know of a present physical object and its contingent properties. Subject uses of 'I' lack recognition, so in such uses the subject cannot refer intentionally to a present physical object.

⁸ I will provide some reasons for holding this claim in chapter 2.

6.2. Can I be aware of a physical object which is present to me without recognising it?

Our question is the following. Suppose we are physical objects with cognitive capacities. Is it true that I have to recognise an object which is present to me in order to know directly of that physical object? Could I be perceptually aware of a physical object without recognising it?

The head-knocking example shows that one can know of a present physical object without recognition. I can know of the object which is me without recognising it. Thus, Wittgenstein's presupposition that recognition is required in order to refer to a physical object by using 'I' was a mistaken one. I can know of the referent of 'I' (even if it is a material object) without recognising this body. This seems plausible, but how is it possible that the subject knows of the object which is the referent of 'I'?

6.3. A mistaken assumption

For Wittgenstein recognition is necessary for knowing of a physical object which is present to me. When, based solely on proprioception, I judge that 'I have my legs crossed', then I cannot misidentify who has her legs crossed. Accordingly, my judgment on this basis is immune. Proprioception is a special way of knowing. Through proprioception the subject cannot know anything but the object which is the subject. Thus there is no need for recognising or singling out the relevant object when an object is known through such special ways of knowing. Thus the lack of recognition in the subject use of 'I' cannot support the claim that 'I' is not a referring expression.

Wittgenstein's line of thought relies on not realising that proprioception is a way of knowing. There are questions about whether proprioception is perceptual, but there is no question whether it is a way of knowing. Proprioception and nociception (for pain) are best understood as perceptual faculties which are directed toward one's own body and its states, which are then present through these perceptual faculties to the subject (Martin 1993 and 1995). Proprioception is a way of knowing even for a Wittgensteinian. According to Anscombe (1972), proprioception is a way of knowing even if it is non-observational and non-perceptual.

I analysed why Wittgenstein thought that recognition is necessary for the referring use of 'I'. I concluded that I can be aware of the physical object which is me without recognising that object or singling it out from among many (the head-knocking case). Thus 'I' can refer to a physical object (if this is the referent of 'I') without that requiring that the subject recognises that object. This is good news. But we face the question of how the subject knows of the referent of 'I'.

7. Summary

For Wittgenstein, subject uses of 'I' which are immune are those which are not based on recognition. In contrast, object uses of 'I' are based on recognition and consequently open to misidentification. According to Wittgenstein, recognition-freedom is necessary and sufficient for the IEM of self-ascription. However, I have shown that it is not sufficient (the mistaken identity example). I provided some reasons to think that IEM requires that we relativise the judgment to its basis (the head-knocking example).

I considered why the IEM of self-ascription could teach us something about the first-person pronoun. For Wittgenstein, the function of recognition is to enable the subject to know of the physical object which is the referent of 'I'. I provided some reasons to think that recognition is not necessary for knowing of a physical object which is present to the subject (e.g. in the head-knocking example). Through proprioception the subject need not recognise the material object which she is in order to know of it and know that it is she herself. Thus, it seems that how the subject knows of the referent of 'I' and the IEM of self-ascription are strongly connected issues.

Chapter 2

Anscombe's Challenge: Guaranteed Right Reference

Abstract

In this chapter, my target is to understand what guaranteed right reference is. Guaranteed right reference (hence GRR) is a double guarantee: (i) against reference failure and (ii) against incorrect reference. Anscombe argues that although in the sensory deprivation tank the subject is able to use 'I', the essential feature of referential uses of 'I' cannot be met. The essential feature is guaranteed right reference.

According to Anscombe's presence constraint, the referent of 'I' has to be perceptually present to the subject when she uses 'I' and this makes GRR possible. I understand this as requiring that the subject has to be acquainted with the referent of 'I'. However, in the sensory deprivation tank, the subject is able to use 'I', she argues, even though the essential feature of referential uses of 'I' (GRR) cannot be satisfied. I will try to work out what the guarantee against incorrect reference amounts to and motivate and reconstruct the presence constraint. This will enable me to pose Anscombe's challenge to the Simple View. If the subject has to be acquainted with herself whenever she uses 'I', then the question arises as to how this is possible in the sensory deprivation tank.

1. Preliminaries

In this chapter, I consider a challenge concerning how the subject knows the referent of her use of 'I' and develop a detailed understanding of GRR inspired by Anscombe. Both Wittgenstein and Anscombe assume that, for uses of 'I' to refer, the subject has to know of the object which is the referent of 'I'. Anscombe, like Wittgenstein, argues against the referentiality of 'I' from the

premise that the user of 'I' cannot know of the referent of 'I' when she uses 'I'. Certain features – GRR, for Anscombe, and the recognition of the referent, for Wittgenstein – would be necessary for referential uses of 'I'. They conclude that uses of 'I' do not refer because the relevant conditions cannot be always satisfied.

There are interesting questions that might be raised concerning why Anscombe thinks that 'I' is not a referring term, particularly why she thinks that the first person pronoun in general fails to qualify as a referring term despite passing standard tests for referentiality. For example, from the first person statement 'I am in pain', we can infer an existential claim 'someone is in pain'. In extensional contexts, we can substitute 'I' in a self-ascription with an appropriate proper name without changing the truth-value of the self-ascription and so on.⁹ But I shall not be engaging Anscombe on these points. Instead, I find it plausible that the first person pronoun refers and that its passing these referentiality tests provides sufficient reason for holding this.

For Anscombe, GRR is the key for understanding how the subject knows of the referent of 'I' – if 'I' refers. GRR is one of the essential features of the use of 'I' in the Reference-fixing Test; the aim of this chapter is to understand GRR and reconstruct Anscombe's challenge.

2. Guaranteed Right Reference

Guaranteed right reference is a double guarantee for a referring expression type. Whenever a token of this type is used (i) **it cannot lack reference** and (ii) **it cannot be used incorrectly or have mistaken reference**. In the literature on first-person reference, GRR is generally assumed to be an essential feature of all uses of 'I' (e.g. Strawson 1994, O'Brien 2007, Peacocke 2008, Campbell 2012).¹⁰ Anscombe describes the phenomenon:

It seems clear that if "I" is a 'referring expression' at all, it has both kinds of guaranteed reference. The object an "I"-user means by it must exist so long

⁹ See Harcourt (2000) and Campbell (2004) for further discussion.

¹⁰ Evans (1982) argues that uses of 'I' lack reference in brain-in-the-vat scenarios. Let me set aside this issue. Evans also argues that 'I' has GRR in the sensory deprivation tank (Evans 1982: 215). This is all I need.

as he is using "I", nor can he take the wrong object to be the object he means by "I". (Anscombe 1975: 151-152).

A guarantee against lack of reference excludes the possibilities that the referent of 'I' is empty or missing. For all tokens of 'I', the existence of the referent in the present is guaranteed. Accordingly, GRR ensures that

GRR1. A use of 'I' cannot lack reference.

It cannot happen that one uses 'I' without successfully referring to an object. Let us call this a **guarantee against lack of reference**.¹¹

Consider, by way of contrast, a term such as 'Homer'. It is possible that 'Homer' was not a name that anybody ever had. It, supposedly, refers to a person. Based on received knowledge, several hundred years ago, tradition might have constructed him from several existing bards and their stories. It is possible that 'Homer' lacks reference. We can try to use a proper name when it lacks reference. In this case, nobody stands in the appropriate (causal) relation to the use of the term. Similarly, one might attempt to use a demonstrative singular term unsuccessfully if there is nothing to be demonstrated or singled out. This cannot happen with 'I' when one uses it. One cannot use 'I' without succeeding in referring to its referent (on the condition that one is not a brain in a vat). Moreover a proper name can be used for an object which does not exist any more, like Cicero. But whenever one uses 'I', the referent exists at the time of use.

There is a second guarantee in GRR, the guarantee against incorrect reference. But it is much less clear what a guarantee against incorrect or mistaken reference means. In Anscombe's paper, we can find two versions of the guarantee against incorrect reference:

A) A guarantee against referring to the incorrect **essential kind of thing**:

¹¹ "But 'I' – if it makes a reference, if, that is its mode of meaning is that it is supposed to make a reference – is secure against reference-failure." (Anscombe 1975: 149).

From Augustine: "The mind knows itself to think", and "it knows its own substance". (Anscombe 1975: 140).

B) A guarantee against referring to **the wrong object**:

The object an "I"-user means by it must exist so long as he is using "I" **nor can he [the subject] take the wrong object to be the object he means by "I"**. (Anscombe 1975: 151-152, with my emphasis).

The first version (A) seems to be implausible. Descartes thinks that we are essentially Cartesian egos, Locke assumes that we are essentially persons and Olson holds that we are essentially human animals. One particular cannot belong to more than one fundamental, determinate essential kind (e.g. lion and human) and has to remain a particular of that same kind until it ceases to exist (Wiggins 1980 and 2001, Snowdon 2014). Accordingly, we have to have one fundamental essential kind. The three options canvassed are mutually incompatible (supposing that our three philosophers are not angels – an angel is its own idiosyncratic kind). Thus, at least two of them are wrong. Descartes might respond that an ideal thinker has to know which kind of thing she is. But, then, this is not a guarantee at all for us in knowing which kind we are, because nothing guarantees that we are ideal thinkers. Consequently, it is plausible that the subject can be wrong about which essential kind she falls under.¹² So if GRR is supposed to hold, then the claim has to be that the subject cannot try to refer to the wrong object by using 'I'. Consequently, the second version looks more promising. This ensures that the subject always uses 'I' for the right object, which is its referent. The subject cannot mean to pick out another object or try to use 'I' for an object other than its referent. We may conclude that we should understand the second guarantee as follows:

GRR2. The subject cannot intend to use 'I' for the wrong object.

¹² For more detailed discussion see O'Brien (1994 and 2007).

I cannot *intend* to refer to any other object but the one which is me by using 'I'; thus I always intend to refer to the right object. Let us call this a **guarantee against incorrect reference**.

In contrast I can intend to refer to the wrong object by using 'KO'. Even if I am KO, I might not know this, and, by using 'KO', I may intend to refer to someone other than me. I simply think it is someone else's name. This cannot happen when I use 'I'.

3. Guarantee against incorrect reference (GRR2)

Let me elaborate what kind of mistake the guarantee against incorrect reference excludes. I cannot intend to refer to the wrong object, if 'I' refers. This should be understood as claiming that the intended referent (what the user intends to refer to) and the semantic referent (what the semantic referent of the referring term in public language is) cannot come apart for the first person pronoun. I will draw on Kripke's (1977) distinction between speaker's reference and semantic reference to explicate Anscombe's notion.¹³

3.1. Speaker's reference and semantic reference

Kripke (1977) differentiates between speaker's reference (which is dependent on the intention of the speaker) and semantic reference (which is dependent on whatever semantic rule we have for determining the referent).¹⁴ The subject might intend to refer to an object by using a singular term when the semantic referent of this term is different from the intended one. In this case, there is a gap between the speaker's referent and the semantic referent of the referring expression; they are different. The speaker takes an object which is not the referent to be the referent of the referring expression. An example will help us to see the difference between the two notions of reference.

¹³ Strawson pushes a similar thought: "I begin with a point which seems to be generally agreed. If someone has and formulates to himself or herself, such a thought as 'I am feeling terrible', then his or her use of 'I' is guaranteed against two kinds of failure to which the uses of some other definite referring expressions are sometimes exposed: it is guaranteed against lack of reference, and it is guaranteed against mistaken or incorrect reference (i.e. against lack of coincidence between the intended reference and the reference conventionally carried, in the circumstances, by the expression used)." (Strawson 1980/1994: 210).

¹⁴ Kripke is responding to Donnellan (1966) and both restrict their discussion to definite descriptions. I try to draw attention to a more general phenomenon.

Call the following case **the coffee affair**. Suppose we are busy waiters and waitresses. There is a man who has not yet paid for his coffee but is preparing to leave. He looks just like John. I remark to you: 'John has not paid yet.' My judgement is based on evidence that someone has not paid yet and my belief that he is John. I intend to refer to the one who has not paid yet. If he is not John, I do not want John to pay for that man's coffee. I want to convey a thought about the one whom I see and whom I think is John. If you know he is not John and do not recognise the similarity, then I fail to convey my thought. The speaker's reference is what the speaker intends to refer to, that object for which the referring expression is used. In our case it is the man who has not yet paid for his coffee, but the semantic reference is John. So speaker's reference and semantic reference can come apart for proper names like 'John'.

3.2. No gap between speaker's reference and semantic reference

I will now apply Kripke's distinction to explicate the guarantee against incorrect reference. When the speaker's reference differs from the semantic reference, then the user intends to refer to something other (that man) than the referent of the referring term (John). And when the speaker's reference and the semantic reference cannot come apart, then the user cannot be mistaken about which object she refers to.¹⁵ In such cases, the user (speaker) cannot intend to refer to the wrong object (i.e. something other than the semantic reference) by using the referring expression. When the speaker cannot intend to refer to the wrong object by using 'I', then the user (speaker) cannot take the wrong object to be the object he means by 'I' – as Anscombe noted. This is not possible unless the user knows which object is the referent of 'I' in a way that he cannot be mistaken about it. This excludes the gap between speaker's and semantic referent. By using 'I', the user will always intend to refer to the right object.

Drawing on Kripke's distinction, we can say that the best way to understand Anscombe's notion of guarantee against incorrect reference is that the speaker's reference and the semantic reference of 'I' cannot be different.

¹⁵ Kripke noted that 'speaker' is not the best word to use, because it can be the writer or the thinker as well whose intentions matter for the notion of speaker's referent.

3.3. Aboutness error

Let me call the mistake when the intended and the semantic reference is different **aboutness error**. To be certain that aboutness error is different from misidentification, we need a clear example involving aboutness error without misidentification.

Let us consider an example where aboutness error is involved but misidentification is not possible. Suppose I know Tony is very clumsy and reliably treads on the toes of people around him. I see a man who I think is Tony, and say 'he will tread on your toes' based on memory and vision. ('He' is used as a demonstrative here.) That man is the only man around besides the two of us. I want to say something about Tony in this situation based on what I know about Tony, so I intend to refer to him. But the one who I think is Tony is not Tony. The semantic referent of 'he' and the intended referent (Tony) are different. Is the judgement immune relative to 'he'? It cannot be misidentified *who is going to tread on your toes* on this basis, yet it is grounded on an aboutness error. This is because I only have reason to think that Tony, but not anyone else, would tread on your toes if he were around you. But the one who is around you is not Tony, so nobody presents the danger of treading on your toes. I am mistaken about whom I see but I cannot misidentify who will tread on your toes. This case shows that aboutness error does not always open the possibility of misidentification. Consequently, aboutness error is a mistake that is independent from misidentification.

To sum up, it seems aboutness-error freedom is behind the guarantee against incorrect reference. I cannot intend to refer to the wrong object by using 'I' because aboutness error is not possible for uses of 'I'. Aboutness-error freedom excludes incongruence between the speaker's referent and the semantic referent.

3.4. Misidentification objection

Perhaps someone might want to challenge my claim that the semantic and intended referent cannot come apart for 'I' with the following case. I can look at a picture and say 'I was 10 years old here'. But the picture is a picture of someone else and I am mistaken. Call this **the photograph case**. Here one

might say that I intend to refer to someone else. It seems that the speaker's reference, the person whom the speaker is speaking of, is the person pictured on the photograph. However, the semantic reference of my use of 'I' is me. So it seems that we have a counterexample to my claim about a guarantee against incorrect reference.

How should I respond to this? The speaker's referent is the person the speaker intends to refer to. The question is: which person do I intend to refer to, which person do I mean when I use 'I'? About whom do I intend to say something? It seems that I intend to refer to myself, and I do say something about myself. But I was mistaken about who is on the picture. It is not a case where I intended to refer to someone other than me. So it is not a case where the semantic referent and intended referent of 'I' come apart.

3.5. Objection

There is a trickier case which might be used against my proposal. Suppose we are searching for someone with messy hair. We are at a super advanced hairdressing class where we try to replicate the messiest hair which we can find. I look at the mirror and say happily 'I have the messiest hair'. It might seem that when the reflection I see in the mirror is not mine, then I would search for someone else.

Does this give us a case where the intended referent and semantic referent of 'I' come apart? No, it does not. I use 'I' because I want to say something about myself. I say, and mean to say, something about me. I am not thinking about that person in the mirror independently of whether that one is me. Although I misidentify who is F, I do not misidentify of whom I am thinking about, to whom I intend to refer to. In contrast, in the coffee affair I want to say something about that man who has not yet paid, independently of whether he is John. I want to say something about him, the one who has forgotten to pay. However, in the messy hair case, **I cannot mean that her hair is messy by saying 'my hair is messy'**. I can use 'John' for someone other than John, but I cannot use 'I' for someone else than me.

I want to suggest that Anscombe is correct in thinking that I need not know which particular object is the referent for using a proper name, but, for using 'I',

I have to know which particular object is the referent without the possibility of mistake. In contrast, it is not a condition on using 'John' that I am not mistaken completely about who I use it for on a particular occasion (when I seem to recognise him).

It is not possible that the intended referent is something other than the semantic referent. What is excluded is only this: that I do not intend to refer to the object which I am. There is an epistemic contrast between how the subject uses 'I' and other singular referring terms – I suggest the asymmetry lies in the impossibility of aboutness error.

3.6. No worry, doppelgangers are safe

We can bring out this epistemic contrast between how the subject uses 'I' and other singular referring terms by considering the following scenario. A doppelganger, an impostor or an indistinguishable twin of John's would look just like John. And even if John were present, I would not know which one is John. I would not be able to tell. If I had an indistinguishable twin or an impostor duplicate or a doppelganger, I would have no problem knowing which one I am.¹⁶ I would never try to use 'I' to intend to refer to my impostor or doppelganger.

In contrast, I can use 'John' when I do not mean to speak about John but his doppelganger, but I cannot use 'I' when I do not mean to speak about myself. If I use 'I', then I mean to say something about that object which is me and I use 'I' for this particular purpose.

The explanation for this is that I know what is the referent of 'I' in a way where I cannot be mistaken about it, but I never know the referent of other referring terms, e.g. 'John', in such a secure way.

4. The Sensory Deprivation Tank Argument and the Presence Constraint

Anscombe's suggestion is that the object always has to be present to the subject whenever the subject uses 'I'. Call this **the presence constraint**. This may well be right. Let me proceed by trying to understand what Anscombe means by

¹⁶ This argument does not apply to the German romantic literature which finds doppelgangers a threat to identity.

‘presence’. According to Anscombe, GRR requires that the referent of ‘I’ is perceptually present to the subject. For Anscombe this means that, whenever I use ‘I’, I am perceptually aware of the physical object which I am. I intend to refer to that object. Just thinking about myself would not guarantee that I intend to refer to the right object. I have to know of that object and the perceptual presence of that object enables me to have this knowledge.¹⁷ This obviously cannot be met in the sensory deprivation tank. My living body is not perceptually present to me.

Let me go back for a moment to our discussion of Wittgenstein. It seems that the idea that the referent of ‘I’ (a physical object) has to be perceptually present to the user of ‘I’ is shared between Wittgenstein and Anscombe. Recall that recognition in Wittgenstein’s non-referentiality argument provides the object which is the referent of ‘I’ and enables the subject to intend to refer to it. Both Wittgenstein and Anscombe use an essential feature of the referential use of ‘I’ to show that this cannot be satisfied for all uses of ‘I’. The only difference is this: Wittgenstein uses recognition and Anscombe uses GRR. It is not possible that they mean the same thing; recognition-freedom is a feature of only some uses of ‘I’, while GRR is a feature of all uses of ‘I’. Accordingly, Wittgenstein’s conclusion is only that some uses of ‘I’ do not refer, while Anscombe’s conclusion is that no use of ‘I’ refers.

Let me reconstruct Anscombe’s non-referentiality argument. This will enable me to arrive to an account of the presence condition and, based on that, I can set up Anscombe’s challenge.

4.1. The reconstruction of the sensory deprivation tank argument

My reconstruction of the sensory deprivation tank argument is indebted to O’Brien’s (2007) reconstruction. The major difference is that I try to work out how the presence constrain relates to GRR in more detail.¹⁸ However my aim

¹⁷ “Just thinking ‘I...’ guarantees not only the existence but the presence of its referent. It guarantees the existence *because* it guarantees the presence, which is presence to consciousness. But NB, here ‘presence to consciousness’ means physical or real presence, not just that one is thinking of the thing. For, if the thinking did not guarantee the presence, the existence of the referent could be doubted.” (Anscombe 1975: 149).

¹⁸ O’Brien’s reconstruction is as follows:

with the reconstruction is different. I try to work out Anscombe's challenge, on the assumption that GRR and our presence condition hold for all uses of 'I', while assuming that uses of 'I' refer, contrary to Anscombe. I start with the argument and will gradually refine it.

Here is Anscombe's non-referentiality argument:

AP1) If 'I' is a referring expression, then every use of 'I' has to have GRR.

AP2) If 'I' refers, then it refers to a material object: a living body.

AP3) The subject has to use 'I' for this object (the living body) if 'I' refers.

AP4) For the subject, the only way to know of the referent of 'I' if it is a material object without the possibility of mistake (satisfying GRR) is that this object is physically present to the subject (in a way that requires direct acquaintance) whenever the subject uses 'I' for it.

AP5) In the sensory deprivation tank the subject can use 'I'.

AP6) In the sensory deprivation tank, the living body, which is the subject, cannot be physically present to the subject (direct acquaintance is missing).¹⁹

Thus, 'I' is not a referring term.

Anscombe supposes that every use of 'I' has GRR if 'I' refers (AP1). For Anscombe, GRR is only possible because the object which is the subject is physically present to the subject – when the subject uses 'I' (AP2-AP4). But, in the sensory deprivation tank, the subject can use 'I' (AP5), yet the subject cannot be acquainted with the object which she is (AP6). AP4 looks to be the crucial premise. Is there any reason to hold it?

1. If 'I' is a referring term, it has guaranteed reference.

2. If the referent of 'I' were not 'physically or really present' to my consciousness, 'I' would not be guaranteed to refer.

3. 'I' is a referring term.

4. So, the referent of 'I' is physically or really present to me.

5. In a sensory deprivation tank I am still capable of self-conscious thought, i.e. if 'I' refers, 'I' refers in a sensory deprivation tank.

6. In a sensory deprivation tank the object meant by 'this body, this human being' is not really present to me. The only thing that could be really present to me is a Cartesian Ego or a stretch of one.

7. So, 'I' refers to a Cartesian Ego, or a stretch of one. (O'Brien 2007: 19).

¹⁹ "If I were in that condition of 'sensory deprivation', I could not have the thought "this object", "this body" – there would be nothing for "this" to latch on to." (Anscombe 1975: 156).

One might find the presence condition unmotivated, but at least some versions of it seem to be shared by others. A version of the presence condition can be found in Wright's work, which I turn to discuss now.

4.2. Crispin Wright and the Presence thesis

Crispin Wright claims that the subject is always present to herself (when she is conscious and awake), and this makes the subject special in the full range of perceivable objects:

It is (merely) that our own presence is, for each of us, a constant factor in the kind of situation, usually but not always social, in which the evidence emerges which bears on various of our psychological characteristics. No-one else is so constantly around us. So no-one else observes as much of us or is as much observed by us. Selves have the best evidence about themselves. (Wright 1998: 101-102).

The quote suggests that the referent of 'I' is always present to the conscious subject and therefore its qualities are present to the subject. Wright speaks about our own presence, and thus everyone's continuous presence to herself when she is awake.

But even if I am always present in some way to myself, why should the use of 'I' be anchored in that presence? I will suggest an answer to this question.

4.3. Acquaintance

I suggest that presence should be understood in terms of direct acquaintance à la Russell. Russell introduced the notion of 'acquaintance' for cases when the subject knows of an object in a direct manner. Russell thought that we are always acquainted with ourselves.²⁰ Acquaintance is used by Russell for acquaintance with sense data. However, Russell (1912) can be read as claiming that 'I' is an exception; we are directly acquainted with ourselves. Accordingly, our knowledge of the object which is the subject is special because it is

²⁰ In contrast to his view we only use acquaintance for acquaintance with a physical object that cannot be a sense-datum.

immediate and direct, deriving from some kind of direct acquaintance with the object which is the subject. Similarly, Kripke assumes that we are acquainted with ourselves by self-acquaintance (2011a: 301).

Campbell characterises acquaintance by epistemic contact with the world and with properties in the world which are not mediated by propositional knowledge.

The idea of 'acquaintance' is the idea of a kind of epistemic contact with a thing or property. (...) Minimally, you might say that what it takes to have epistemic contact with b is this: you must have a piece of propositional knowledge whose content can be specified using a term referring to b, such as 'b'. (...) Russell argued that our knowledge of things cannot, in general, be explained in terms of our knowledge of truths. Russell thought that there were two sorts of knowledge: knowledge of truths, and knowledge of things. (...) Knowledge of truths depends on acquaintance with objects. For it is acquaintance with objects or properties that provides our knowledge of reference[.] (Campbell 2009: 1-3).

Campbell uses Russell to point out that acquaintance with objects enables us to think about and make judgements concerning the objects we are acquainted with.

I want to distinguish acquaintance from direct acquaintance. I use **direct acquaintance** for such cases when the subject directly knows of an object (or a property). We can find two types of acquaintance in Evans's discussion. One is indirect and involves reliable causal chains which are preserved in testimony and provides a relation between the subject and the referent of the term. A reliable causal relation (like a causal chain through testimony) is not sufficient for the subject to be directly acquainted with the relevant object. The object has to be present to the subject; this requires a direct relation such as perception. The object has to be in the vicinity of the subject and the subject has to have a basic way of knowing of that object providing the direct epistemologically useful relation to the object. Such a direct relation is called direct acquaintance.

A basic way of knowing – like seeing – is such that it does not require

another way of knowing. When one sees an object then one knows of that object directly. When I see an object, then I am aware of it through a basic way of knowing, but when I remember seeing that object then my way of knowing is not basic. When one informs me that there is a red top hat in this room where we are, but I neither see nor touch it (nor know about it in another way) then I am not directly acquainted with the red top hat. Memory, inference or testimony are such that they require other ways of knowing (by the subject or by others). They cannot be the original source of knowledge except in rare cases. In contrast, introspection, proprioception, audition and smell are basic ways of knowing. They do not require as preconditions other ways of knowing to provide the content known through them. Direct acquaintance requires a basic way of knowing through which the object or the property can be known for the subject. I propose that the subject has to be directly acquainted with herself whenever she uses 'I'.

4.4. Is Anscombe's argument is any good at all? – O'Brien's criticism

I have followed O'Brien's reconstruction of Anscombe's argument in broad stroke. O'Brien thinks that the argument is flawed. Anscombe holds that the only thing which can be known in the sensory deprivation tank as the object which is the referent of 'I' is a thinking thing like a Cartesian Ego. O'Brien's response is the following: "[...] while it may be admitted that what 'I' refers to must be identified as a thinking thing, that admission is not sufficient to rule out the possibility that what 'I' refers to is also a bodily thing. For we can say that I do identify a body when I use 'I' it is just that I do not realise that that is what I am doing. This point strikes me as fatal to the argument as it stands." (O'Brien 2007: 26). So, according to O'Brien, even if I think of myself as a thinking thing this does not rule out the possibility that what I think about as a thinking thing is a physical object as well. O'Brien's idea is this: I can identify or know of an object without knowing that it is a physical body. I may know of an object without knowing which kind it is. Thus, the argument is flawed. It does not follow that 'I' does not refer.

Despite this, we still can learn something from it. My suggestion, following Anscombe, is this. Direct acquaintance with the object which is the subject

should be sufficient to make a guarantee against incorrect reference possible. Uses of 'I' could have a guarantee against incorrect reference because the subject, whenever she uses 'I', has to be directly acquainted with the object which she is (in a way that excludes the possibility of mistake). This presents a challenge to the view which I will propose.

5. Anscombe's challenge

Anscombe poses a challenge for views such as the Simple View, where I have to be directly aware of the object which is the referent of 'I' whenever I use 'I'. If direct acquaintance (which need not be perceptual) of the referent of 'I' is necessary for using 'I' referentially, then how can the subject use 'I' in the sensory deprivation tank without direct acquaintance with its referent?

Accordingly, the sensory deprivation tank seems to pose a challenge *if we assume that the presence constraint is correct*. In the sensory deprivation tank, a subject can use 'I' but the uses of 'I' cannot have the essential features of the referential uses of 'I'. If uses of 'I' refer then the subject has to be directly acquainted with its referent for all uses of 'I' and 'I' has to have guaranteed right reference. But in the sensory deprivation tank it is not clear how the subject can be directly acquainted with the object which she is, although she is able to use 'I'. This is Anscombe's challenge.

Anscombe's challenge should be answered. I will answer it in the last chapter of this thesis, when I will consider objections to the Simple View.

Chapter 3

The Basic Characterisation of IEM

Abstract

In this chapter, I characterise the basic notion of IEM through explicating the notion of misidentification that is involved in IEM. As I will argue, a judgment, *Fa*, based on basis B, is immune iff the subject cannot misidentify, on this basis, B, that, this object, *a*, is *F*. Accordingly, misidentification happens when the subject thinks *Fa* on a basis, B, which is for thinking *Fb*. Pryor (1999) offers an alternative to this standard notion of misidentification: which object misidentification. I will argue that the standard notion of misidentification does a better job at characterising IEM.

1. Outline

The plan of this chapter is as follows. First, I start with basic examples of misidentification, and, from that, I will explicate the standard notion of misidentification involved in understanding IEM. Second, I will discuss an alternative to this standard notion of misidentification from Pryor: which-object misidentification. Pryor argues that which-object misidentification should replace standard misidentification in our account of IEM. Third, I disagree with Pryor on this point and will argue that there is no reason for this replacement for understanding IEM present tense self-ascriptions.

2. Examples of standard misidentification

Let me turn to my first example of misidentification. Wittgenstein's (1958: 66-67) example of misidentification is as follows: I am wrestling with someone and I see a hand covered with blood, so I judge that 'I am bleeding' based on vision. This judgement is open to misidentification on this basis. The one who is bleeding might be the one with whom I am wrestling. In this case my basis

could have been for the judgement 'He is bleeding', but I judge 'I am bleeding' mistakenly, assuming that I see myself.

Let me change the basis and test whether misidentification is possible. Suppose I am chopping onions and accidentally cut my finger. I close my eyes but I feel that I am bleeding. Only based on feeling, judging 'I am bleeding' will not open the possibility of misidentification. On this basis I cannot think of anybody but me.

Consider a different case. I bleed very easily and I accidentally knock my knee into some sharp furniture. Only based on feeling, I judge that I am bleeding. Here again misidentification is not possible. But if I were looking at a mirror and see that my knee is bleeding, I might judge on this basis, vision, that I am bleeding. But in this case, on this basis, misidentification would be possible.

Another standard example is the mirror case. I might look at the mirror and judge, based on vision involving the mirror, that 'I have messy hair', even though the one whom I see in the mirror is not me. On this basis, vision, I am justified in judging that 'She has messy hair'. It would be a different question if my friend says to me that I have messy hair and based on testimony I judge 'I have messy hair'. Then the question would be whether based on testimony I can misidentify who has messy hair. When I hear "you have messy hair", or "Krisztina has messy hair" then it could be that 'you' or 'Krisztina' refer to somebody other than me. Nothing excludes this possibility so misidentification is possible.

In yet another situation, the basis can be vision without involving a mirror. If I have long hair and I see my hair all over messed around, and judge based on vision that I have messy hair, then, based on vision, I can misidentify who has messy hair. It seems the basis has to be specified and we have to decide whether on this basis misidentification is possible.

Consequently, misidentification happens when a subject on basis B judges *Fb* when this basis B is for judging *Fa*.²¹

²¹ Pryor only discusses circumstantial immunity which is basis-relative.

2.1. Do we have to accept Basis-relativity?

Basis-relativity is widely accepted in the IEM literature (Evans 1982²², Peacocke 2008, 2012, Recanati 2007, 2012, O'Brien 2007, 2012, etc.). I have given some sense of its plausibility above. It would seem plausible to accept that we have to relativise the judgment to its basis, *Fa* (B) and allow that the same judgment is immune on one basis but not on another basis. Thus, IEM is plausibly thought to be basis-relative.

There is good reason to agree with this point. When I judge, based on proprioception, that 'I have my legs crossed' I cannot test whether it is immune based on proprioception and vision. I cannot test the judgement in this way, because based on proprioception and vision the same judgment would be **not** IEM. I would be able to misidentify whose legs are crossed – the one whom I see. Accordingly, what matters for IEM is what is not possible on a particular basis, a way of knowing. This is a good reason to accept basis-relativity.²³

3. The standard notion of misidentification

Misidentification happens when it seems to me that I have a basis for judging *Fa* based on basis B, but, actually, the basis B is for judging *Fb*. I call cases when the subject attributes something to *a* when her basis is actually for attributing to *b*, cases of **ill-grounded singular thought** (hence **ILL**). (If there is no object which the subject is thinking of when the subject tries to form a singular thought, then there is no singular thought. It is not a case of ill-grounded singular thought in my terminology.) And I call cases when the subject judges and is justified in judging something about *a*, cases of **well-grounded singular thought** (hence **WELL**). **WELL** does not exclude that the predication is mistaken. **WELL** only excludes that the thought is of the wrong object.

In a case of standard misidentification, *Fa* is an ill-grounded singular thought (**ILL**) on a particular basis *B*₁, while *Fb* is well-grounded singular thought (**WELL**) on this very basis *B*₁.

²² Evans claims that the judgment 'I have my legs crossed' which is based on proprioception is immune but the same judgement based on vision is not IEM (Evans 1982: 220).

²³ I will argue for basis-relativity in the next chapter. But one could accept basis-relativity only for the sake of argument here if the reasons above do not yet seem sufficient. Pryor has to accept it because he denies absolute IEM; believing in absolute IEM requires the rejection of basis-relativity.

Pryor (1999) developed an account of IEM which departs from the standard account of misidentification sketched above. He claimed that there is a more basic notion of misidentification: which object-misidentification. He contends that which object-misidentification should replace the standard notion of misidentification. I now turn to Pryor's distinction between the two kinds of misidentification.

4. Pryor's *de re* misidentification

Pryor distinguishes between two kinds of IEM by distinguishing between two different kinds of misidentification. The first is the ***de re* misidentification**:

- (i) There is some singular proposition about x , to the effect that it is F , that a subject believes or attempts to express. [$Fx = \text{ILL}$]
- (ii) The subject's justification for believing this singular proposition rests on his justification for believing, of some y and of x , that y is F and that y is identical to x . [$Fy = \text{WELL}$]
- (iii) However, unbeknownst to the subject, $x \neq y$. (Pryor 1999: 274-275; with my parenthetical remarks).

Fx is a singular proposition, thus ' x ' is not an open variable requiring quantification (existential/universal) but refers to a particular object. It functions like Fa in our formulation of the standard view.

So it seems that Pryor's notion of *de re* misidentification requires that the subject judge Fa (B), ILL, when the subject's basis is actually for judging Fb (B), WELL. This is exactly the standard notion of misidentification which I have introduced above.

5. Pryor's which object misidentification

Let me turn to the second kind of misidentification which Pryor discusses: which object-misidentification (hence **wh-misidentification**). Pryor introduces wh-misidentification by an example:

I smell a skunky odor, and see several animals rummaging around in my garden. None of them has the characteristic white stripes of a skunk, but I believe that some skunks lack these stripes. Approaching closer and sniffing, I form the belief, of the smallest of these animals, that it is a skunk in my garden. This belief is mistaken. There are several skunks in my garden, but none of them is the small animal I see. (Pryor 1999: 281)

How is Pryor's example supposed to work? First, I have a basis to believe that 'One or more animals are skunks in my garden' based on smell – an existential thought. I might not form this thought but I am at least in a position to form this thought. Second, I believe that 'that one is one of the skunks/is the skunk in my garden' based on vision and smell – a singular thought. But this is mistaken; this is not one of the several skunks in my garden. Thus I misidentify which object is one of the skunks in my garden. Pryor dubs this kind of misidentification: **Which object-misidentification (wh-misidentification)**.

Here we have an existential and a singular thought. An existential thought, something is *F*, requires only that one or another object is *F*; the object which is *F* need not be specified, discriminated or identified. I might see a crowd of animals and think about them without identifying or singling out one of them. The existential thought is a well-grounded **existential** thought, but in this case (when I see a crowd of animals) it is not a well-grounded **singular** thought. It is not a singular thought at all.

Let me elaborate the difference in the ground for an existential and for a singular thought. I cannot think a singular thought, *Fa*, unless I think about *a*. Singular thought requires thinking about a particular object. In contrast, I can think an existential thought about a mob without thinking or being in a position to think of one particular person in the mob.

Note that wh-misidentification is different from *de re* misidentification in several respects. Wh-misidentification does not involve two singular thoughts and it is not the case that one thinks *Fa* (*B*) instead of *Fb* on the same basis *B*. The subject, *S*, is not in a position to form a singular thought about *y*. *S* can only form an existential thought about some objects. But the subject is not in a position to know of the particular object, *y*, which would be necessary for

forming a singular thought about *y* (Pryor 1999: 282). This is the key difference between *de re* misidentification and wh-misidentification.

5.1. Wh-misidentification should replace *de re* misidentification

Pryor (1999: 285) writes that: “Wh-misidentification can occur without *de re* misidentification [as the skunk case shows].” I agree with this. However, Pryor holds that from immunity to wh-misidentification (hence **wh-IEM**), *de re* IEM follows, thus wh-IEM is more basic and more fundamental than standard *de re* IEM (Pryor 1999: 285-286). I disagree. I will argue that a logical entailment is neither sufficient to make it more fundamental nor to make it more basic for understanding IEM.

Pryor points out that ‘I have a pain’ based on introspection and ‘I think it will rain’ based on introspection are plausibly immune to wh-misidentification (Pryor 1999: 283, 287). So if wh-IEM is more basic (i.e. standard IEM follows from it) then we have to understand IEM by employing wh-misidentification instead of standard misidentification.

5.2. Why does Pryor think that wh-misidentification should replace *de re* misidentification?

Pryor argues that immunity from wh-misidentification is what should figure in our understanding of the phenomenon of IEM and not immunity from standard *de re* misidentification. “[I]f a belief is immune to wh-misidentification when justified by certain grounds, that entails that it is also immune to *de re* misidentification when justified by those grounds. (...) [I]mmunity to wh-misidentification entails, but is not entailed by, immunity to *de re* misidentification. In that sense, immunity to wh-misidentification is a more basic and more rare epistemic status.” (Pryor 1999: 285-286).

Let me examine the argument for thinking that wh-immunity is more basic than *de re* immunity. Pryor argues as follows: (i) from wh-immunity standard *de re* immunity follows, but not vice versa and (ii) paradigmatic examples of IEM (like ‘I have a pain’ judged on the basis of introspection) have wh-immunity. Thus, wh-IEM is the more basic notion.

6. Why wh-IEM instead of de re IEM?

To understand Pryor we need to understand why he would hold that wh-misidentification is the most basic and strongest kind of misidentification for understanding IEM. There are two reasons. The first is that from wh-IEM *de re* IEM follows (Pryor 1999: 285). I will show that this is trivially true and cannot be the real reason.

Wh-misidentification happens when S judges *Fa* when S is actually only in a position to judge 'Something is *F*'. If the judgment *Fa* is wh-immune then it is not possible that S is in a position to judge 'Something is *F*' when *Fa* turns out to be false. 'Something is *F*' simply follows from *Fa*. What enables me to judge *Fa* enables me to judge 'Something is *F*'. This is true; but this is trivial. So it is trivial that S cannot be in a position to judge *Fa* when S cannot judge 'Something is *F*'. Consequently, this alone cannot suffice to show that it provides a more fundamental understanding of IEM. The subject matter of IEM seems to be an epistemological one (the subject's immunity to a certain kind of error) and not a logical one.

From 'This is a chair' it follows 'The world exists'. This is a chair cannot be true unless the world exists, and presuppositions follow from what presupposes them. 'This is a chair' is immune relative to 'this'. But does it follow that 'The world exists' is more basic or more fundamental for the question of immunity of this judgement? Is it true for all presuppositions that they are more basic? Even if this were true, they are not more fundamental in understanding IEM. The question whether the world exists is usually thought to be more fundamental than whether this is a chair, but this question is orthogonal to the question of IEM of 'this is a chair' relative to 'this'. Similarly, it is not clear that 'p' is more basic or more fundamental than 'p and q' because the former follows from the latter. In the case of 'p and q', 'p' is a component of the former; perhaps this makes it more basic. But 'Something is *F*' is only a presupposition of *Fa*, and is not a component of it, so this reason is unavailable.

Pryor might argue that when I cannot judge that a particular object *a* or another particular object *b* is *F*, then I might still be able to judge 'something is *F*'. His thought is that when I cannot judge the existential thought then this

provides a stronger kind of immunity than *de re* IEM. But there is a question whether this gets the subject matter right.

Immunity was introduced in understanding singular thought about a particular object. A singular thought is about a particular object which the subject knows of in one or another way. It is not a thought about one or another object or many objects, like existential thought. In contrast, an existential thought is about one or another object or many objects. It could be about an object of which the subject is acquainted with, but this need not be the case. From '*Fa*' it follows that 'Something is *F*'. But without being able to think about a particular object, one can have grounds to think about many objects. When I see a crowd then I see that crowd even if I cannot identify or reidentify any of its members. To think of a particular object is a different phenomenon from thinking of many objects without any one in particular. Thus, whether I can think an *existential thought without a singular thought* (There are skunks in my garden) or whether I can think a *singular thought* about an object (This is one of the skunks in my garden/ I have messy hair) would seem to be two different questions.

Pryor wants to rule out wh-misidentification, the ability to think 'Something is *F*' after being defeated in thinking of *a* that it is *F*. The thought that we should rule out the possibility that I can think an existential thought after being defeated in thinking a singular thought is not a very natural one. If one wants to argue that existential thoughts are more fundamental than singular thoughts then one should argue for this claim directly. In that case, one should argue that one has to either have a ground to think an existential thought and, derived from this, a ground to think a singular thought, or that for every singular thought, the subject has to have a general ground as well. Neither of these options seems plausible. Thus it is not clear why an existential thought must be relevant in thinking a singular thought. So before one argues for the fundamentality of existential thought over singular thought, it seems that the question whether an existential thought is possible is orthogonal to whether a mistake with a singular thought is possible. Unless one argues that existential thoughts are more fundamental, the default position would seem to be that

existential and singular thoughts are simply different, and questions about one cannot be reduced to the question about the other.²⁴

Suppose I think I am in a position to think Fa , but if I am only in a position to think 'Something is F ', then I was mistaken in thinking that I have a ground to think a singular thought. But this mistake is a different mistake from misidentification. It is a mistake about which kind of ground I have, whether I have grounds to think a singular thought or only an existential thought (without a singular thought). When I see someone as a person or recognise him, then I have a singular ground to think about him. When I see a crowd of people and hear loud voices, I have grounds to think existential thoughts about people exemplifying certain qualities. When I cannot discriminate anybody (neither by vision nor by audition or by touch) in the crowd then I cannot form a singular thought about a member of the crowd.

It is not clear why we need an existential ground for a singular thought. When the ground for the singular thought is defeated, it is unclear why there must be an existential ground that would remain untouched. An existential thought 'Something is F ' follows from Fa/Fb . So if my singular ground for Fa/Fb is defeated, then the existential ground which follows from it is thereby defeated as well. So it is not clear how an existential ground would remain untouched after the singular ground is defeated. A case when an existential ground could be undefeated by the defeat of the singular ground is when one has an independent existential ground. If there is an existential ground it should be completely independent from the singular ground, like in the skunk example. I have a vision-based singular ground to think about the animal, which I see, and a vision and smell-based general ground to think about the animals in my garden. My ground for thinking about a particular animal cannot be derived from a ground for thinking about many animals. They are different kinds of grounds. But if a subject, S , has a singular ground to think a singular thought Fb ,

²⁴ It seems that a Russellian analysis of singular thought requires the reduction of a singular thought to an existential thought. So, for Russell, there would appear to be reason to think that existential thought is more basic and more fundamental than singular thought. But the project of understanding singular thoughts (from Evans and others) was initiated to overcome this. The IEM literature tries to answer how singular thought is possible without reducing it to existential thought. So it would seem that if singular thought is reducible to existential thought in a Russellian manner then we would not have genuine singular thought. IEM is about how genuinely singular thoughts are possible.

but S thinks *Fa*, then this is a case of standard *de re* misidentification. Therefore it is unclear what reason there could be for thinking that wh-misidentification is more basic and more fundamental.

To repeat, the main objection against wh-misidentification is this: if I have no reason to think about many objects, then the fact that I cannot think ‘Something is *F*’ is not surprising. But if I have reason to think about many objects without having reason to think about a singular object, then it is not clear why I would think a singular thought in the first place.

6.1. Does Pryor’s skunk example satisfy basis-relativity?

I return now to Pryor’s key example concerning skunks. Does Pryor’s skunk example satisfy basis-relativity? For misidentification when I think *Fa* based on B then on this basis I can misidentify which object (from many) is *F*. Recall, first I have a basis to believe that ‘One or more animals are skunks in my garden’. The basis is smell. Second, I believe that ‘that one is one of the skunks/is the skunk in my garden’ based on vision and smell. This has an additional basis: vision. But that one is not one of the skunks in my garden. Thus I misidentify which object is one of the skunks in my garden.

It seems that the example cannot work for a judgment relativized to a basis like vision and smell. But this might mean only that the example has some problem that can be easily fixed. Do we face a deep or a superficial difficulty? The example can probably be reconstructed in a way where both the existential thought and the singular thought are based on vision and smell.

I think the real difficulty comes from the fact that there is an independent existential ground which remains untouched by the defeating evidence against the singular thought. And whether or not there is such an independent existential ground does not seem to challenge whether misidentification is possible.

It seems that we do not find any reason to think that an existential ground which is independent from the singular ground should play any role in the characterisation of IEM. There might be wh-IEM and it could be interesting with respect to another epistemological question, but there is no reason to characterise IEM with wh-misidentification.

7. Summary

In this chapter, I provided the basic characterisation of IEM. I proceeded by explicating the notion of misidentification that is involved in IEM. According to this, a judgment *Fa* based on basis B_1 is immune relative to 'a' iff the subject cannot misidentify **on this basis B_1** which particular object is *F*. Hence, misidentification happens when the subject thinks *Fa* on a basis, B_1 , which is for thinking *Fb*. Pryor (1999) offers an alternative to this standard notion of misidentification: which object misidentification. I argued that the standard notion of misidentification does a better job at characterising IEM.

Appendix: Departing from Pryor

Pryor could have another reason for using wh-IEM: the explanation of the IEM of memory judgements may require wh-misidentification. This reading is supported by the fact that Pryor concentrates on memory-based judgements, and wh-misidentification seems to have been introduced for answering the question whether quasi-memory judgements are immune. I will argue that memory-judgements are different from present tense self-ascriptions. It is not clear that whatever is true of memory-based judgements will be true of present tense self-ascriptions. But first I have to discuss quasi-memory.

Quasi-memory

Shoemaker introduced quasi-memory judgments for cases involving fission. Suppose Tweedle undergoes fission, so two people after fission, Dum and Dee, quasi-remember what happened with Tweedle prior to fission. (Dum and Dee have the same 'memories' – which originated from Tweedle.) Dum quasi-remembers that he pulled a cat's tail. But, for Shoemaker, judgements based on quasi-memory will not be immune to standard *de re* misidentification. This is because it can turn out that what Dum quasi-remembers, pulling the cat's tail, is someone else's memory. In our case it is Tweedle's memory and not Dum's. Pryor attempts to show that Dum's and Dee's quasi-remembering will not be

immune to wh-misidentification, but will be immune to standard *de re* misidentification. For Pryor, quasi-memory is not immune to wh-misidentification.

Memory-based judgements are not the basic case for IEM

I will argue that memory-based judgements cannot be paradigmatic examples of IEM. Memory-based self-ascriptions have a different kind of IEM from the IEM of present-tense self-ascriptions. I will list some obvious reasons for accepting this difference.

First of all, the IEM of present tense self-ascription made on a certain basis is such that misidentification is excluded only for the subject and from the subject's point of view. (When I say 'I am in pain' you can misidentify who is in pain if you think someone else is speaking.) Memory-based judgements could be immune only if memory is factive. Memory is factive if I cannot remember something unless it happened. However, I can seem to remember that I was in Paris because your description of your Paris trip was so vivid that I seem to remember that it happened with me. Even if it is true that I only seem to remember, from the subject's point of view misidentification **can** happen. The question for 'I am in pain' based on feeling, is whether misidentification can or cannot happen from the first person point of view. So if memory is factive but I might not know whether it counts as a memory when I seem to remember, then memory-based self-ascriptions are very different from present tense self-ascriptions in this respect. I seem to remember the experience of being there. So even if memory-based judgements can be immune, they cannot be so from the first person perspective. In contrast, IEM of present tense self-ascription is only from the first person perspective.

Second, memory is not a basic way of knowing. My memory only preserves the content which I have gained through another way of knowing like vision or touch. When I remember seeing or touching something then the original way of knowing is vision or touch. A basic way of knowing cannot require another basic way of knowing as the original source. Vision or touch are basic ways of knowing; they do not require other ways of knowing. They alone could put the subject in a position to know something when one perceives something.

Testimony (similarly to inference²⁵) requires another way of knowing so it is not a basic way of knowing.

We might have to decide case by case. Remembering that I was in Paris requires other ways of knowing which were employed in gaining the experience which I remember. When only visual experience is involved then vision is the prior source. If food is involved in the experience of being in Paris then vision, smell and taste are the prior ways of knowing.

The IEM of present tense self-ascriptions made on the basis of a basic way of knowing seems to be different from memory-based self-ascriptions where content from a basic way of knowing may be preserved or distorted in memory. Thus we have to concentrate on the basic case.

²⁵ Inference usually requires premises, and premises require further ways of knowing, or at least in most cases. Inference might be a basic way of knowing in some special cases, when all premises are presuppositions of everything what we know. E.g. I implicitly take it to be the case that the world around me exists. What follows from this does not rest on any other basic way of knowing. In this case inference could count as a basic way of knowing.

Chapter 4

The Myth of Absolute Immunity

Abstract

I will argue against Shoemaker's distinction between absolute and circumstantial IEM. In particular, I will question the possibility of absolute IEM. For Shoemaker, absolute IEM is different from circumstantial IEM, because an absolutely immune self-ascription is immune in all circumstances, while a circumstantially immune self-ascription is immune in at least one circumstance, but not in all circumstances. An absolutely immune self-ascription is immune on all bases. Accordingly, absolute IEM is basis-independent. This denies basis-relativity, where a self-ascription is only immune relativized to certain bases, but not to all bases. I will argue for basis-relativity by using counter-examples and by providing a general argument why absolute IEM (basis-independence) is unattainable.

1. Preliminaries

In 'Self-Reference and Self-Awareness' (1968), Shoemaker identified a special class of first person judgments that exhibit a certain kind of property. He labelled this class as the class of judgments that are *immune to error through misidentification* relative to 'I' and discussed it extensively. Shoemaker credits Wittgenstein with first introducing the phenomenon. Like Wittgenstein, Shoemaker concentrates on understanding the IEM of self-ascriptions. However, his account differs from Wittgenstein's. Shoemaker differentiates between two kinds of IEM – absolute and circumstantial IEM – and accepts that 'I' is a singular referring term that picks out a material object. It is obvious that for understanding immunity it is necessary to reconstruct Shoemaker's

treatment of IEM and critically engage with it.²⁶ Consequently, I need to examine whether there are two kinds of IEM as Shoemaker suggests.

My strategy will be to argue against the possibility of absolute IEM. An absolutely immune self-ascription is immune on all bases. So absolute IEM denies basis-relativity (where a self-ascription is only immune relativized to certain bases but not to all bases). My main aim is to argue that IEM is basis-relative; this means that all kinds of IEM are basis-relative. This excludes the possibility of absolute IEM. In order to show this, I will first provide some counterexamples against absolute IEM and then offer a general argument against the possibility of absolute IEM.

Although there seems to be a consensus in the immunity literature that IEM is basis-relative there is no argument presented directly for this claim.²⁷ Evans, for example, claims on the basis of one example that a judgment is immune on one basis, but not immune on another basis (1982: 218-224). This seems to be accepted as sufficient reason for believing that IEM is basis-relative for all self-ascriptions. I find this plausible, but from a small set of examples a universal generalisation does not follow. The universal generalisation is that all cases of immune self-ascriptions are basis-relative. The truth of this claim is crucial in my argument for the Simple View, so I have to argue for this conclusion.

In this chapter, let me restrict the target notion of IEM to the IEM of self-ascriptions. Whenever I talk of IEM, I mean the IEM we find in putatively immune self-ascription.

2. Shoemaker's Picture of Immunity

It is only natural to start with an overview of Shoemaker's picture of immunity. Shoemaker holds that a judgment like 'I am in pain' is immune relative to 'I' because it cannot be the case that I am wrong about who is in pain, but right that someone else is in pain:

²⁶ Even though Shoemaker holds that demonstrative judgments can be immune, he thinks that this kind of immunity is different from the immunity of self-ascriptions (Shoemaker 1968: 558-559). I agree, and will argue for this claim in chapter 6.

²⁷ Higginbotham (2012) might be an exception if we read his notion of 'logical immunity' as absolute immunity. Settling this issue is out of the scope of my discussion. Another exception is Longuenesse (2012); I will discuss her explanation of IEM in chapter 10.

The statement "I feel pain" is not subject to error through misidentification relative to 'I': it cannot happen that I am mistaken in saying "I feel pain" because, although I do know of someone that feels pain, I am mistaken in thinking that person [who feels pain] to be myself." (Shoemaker 1968: 557).

If I say "I am bleeding," it can happen that what I say is false even though I am giving expression to the knowledge that a certain person is bleeding; it may be that I do see a bleeding arm or leg, but that because my body is tangled up with that of someone else (e.g., we are wrestling) or because I am seeing my identical twin or double in a mirror, I am mistaken in thinking the person who is bleeding to be myself. (Shoemaker 1968: 556).

This is in perfect accordance with the basic notion of misidentification (as described in the last chapter) except that the self-ascription of being in pain is supposed to be immune on all bases.

Shoemaker distinguishes between two kinds of IEM. I now turn to examine what marks their difference.

3. Absolute and Circumstantial Immunity

We need a preliminary grasp of the two notions: 'absolute immunity' and 'circumstantial immunity'.²⁸ An 'a'-judgment (a judgment concerning 'a', like *Fa*), *Fa*, is **absolutely immune** relative to 'a' iff who is *F* cannot be misidentified in any circumstance. This requires that no change even in the basis of *Fa* can affect its immunity. Thus, it follows from being absolutely immune that it has to be immune on all bases. In contrast, a 'b'-judgment, *Fb*, is **circumstantially immune** iff who is *F* cannot be misidentified in at least one circumstance, but not all circumstances are like that. The relevant feature individuating the circumstance might be the basis on which the 'b'-judgment is made.

Shoemaker introduces the notion of 'circumstance' through contrasting two cases. In the first circumstance, based on seeing a table, I judge that *I am facing*

²⁸ Shoemaker distinguishes between two further kinds of immune judgements: logical and factual immunity (Shoemaker 1984b). These differ from absolute and circumstantial immunity and pertain primarily to memory and quasi-memory judgments. I will not discuss them here, since my focus is present tense self-ascription.

a table. In the second circumstance, based on seeing myself in the mirror facing a table, I judge that *I am facing a table* (Shoemaker 1968: 557). I only take it that if the basis (in Shoemaker's sense) is different, then the circumstance is different – which would be true on all readings of 'circumstance' (even if something else could also make a difference on some readings of it).²⁹ In order to argue against absolute IEM it is sufficient to show that, for a supposedly absolutely immune self-ascription, misidentification is possible on one basis.

Absolute IEM requires that a self-ascription which is absolutely immune cannot be misidentified relative to 'I' on any basis in all circumstances. In the previous chapters, we have found some reasons to believe that immunity is basis-relative. Absolute IEM and basis-relativity are incompatible; the former denies that one kind of IEM (absolute IEM) is basis-relative.

3.1. Shoemaker's Two Theses

I will reconstruct the basic idea behind absolute and circumstantial IEM in the form of two theses, based on quotes from 'Self-reference and Self-awareness'.

3.1.1. The Circumstance-independence Thesis

According to the first thesis an absolutely immune judgment is immune in all circumstances, so there can be no circumstance where the subject can misidentify who is *F*:

First-person statements that are immune to error through misidentification in the sense just defined, those in which 'I' is used "as subject," could be said to have "absolute immunity" to error through misidentification. A statement like "I am facing a table" does not have this sort of immunity, for we can imagine circumstances in which someone might make this statement on the basis of having misidentified someone else (e.g., the person he sees in a mirror) as himself. But there will be no possibility of

²⁹ Peacocke (2008: 92-103) uses the notion of circumstance for differentiating between normal and abnormal circumstances. Similarly, Evans (1982: 186) speaks about normal ways of gaining knowledge. But this is not Shoemaker's point in keeping the notion of circumstance in play; he never contrasts normal with abnormal circumstance. 'Circumstance' could mean circumstance of evaluation (Kaplan 1989) but nothing seems to support this claim either.

such a misidentification if one makes this statement on the basis of seeing a table in front of one in the ordinary way (without aid of mirrors, etc.); let us say that when made in this way the statement has "circumstantial immunity" to error through misidentification relative to 'I'. (Shoemaker 1968: 557 with my emphases)

Accordingly, the first thesis ensures that absolutely immune self-ascriptions are immune independently of the circumstance in which the self-ascription is made. Call this the **circumstance-independence thesis**:

Circumstance-independence Thesis: *A self-ascription, F_i , is absolutely immune iff the subject cannot misidentify who is F in any circumstance.*

Correlatively we can introduce the notion of circumstantial immunity:

Circumstantial IEM: *A self-ascription, F_i , is circumstantially immune iff the subject cannot misidentify who is F in at least one circumstance, but not all circumstances are such.*

If, for F_i on one basis, the subject can misidentify who is F , then F_i cannot be absolutely immune. To understand what is required we have to know what differentiates one circumstance from the other. In the text it seems that, for circumstantial immunity, who is F can be misidentified on one basis (misidentification in the mirror), but on another basis (direct vision), it cannot be misidentified.

3.1.2. The Basis-independence Thesis

It is an interesting question how a circumstance is individuated. As I noted earlier, Shoemaker only provides a contrast between two cases. However circumstance should be individuated at least by the difference in the basis of the self-ascription. This much is clear. Absolutely immune self-ascriptions will be immune to error through misidentification relative to 'I' in all circumstances. Therefore they are immune on all possible bases. If I deny basis-independence I

would thereby deny circumstance-independence. Therefore the claim I target is:

Basis-independence Thesis: *For an absolutely immune self-ascription, Fi (B), the subject cannot misidentify who is F , whatever the basis.*

On all bases on which a self-ascription Fi can be asserted, Fi will be immune. Absolute immunity makes the basis irrelevant in the sense that the status of being absolutely immune does not vary depending on which basis the judgment is made.

If absolute IEM is basis-dependent, then it cannot be true that all kinds of IEM are basis-relative. Thus, the basis-independence of absolute IEM denies that all kinds of IEM are basis-relative. In contrast, circumstantial immunity is basis-dependent. A circumstantially immune self-ascription is immune only on some but not on all bases (circumstances). A change in the basis of the self-ascription can affect its immunity. Thus Shoemaker accepts that circumstantial IEM is basis-dependent, but denies that absolute IEM is basis-dependent.

We now have a basic understanding of absolute and circumstantial IEM. So I will turn to argue for basis-relativity by denying the possibility of absolute IEM. For an absolutely immune self-ascription based on basis B_1 a change in the basis (B_2) of the self-ascription cannot affect its immunity. If absolute IEM requires that a change in the basis cannot affect its IEM, then in order to provide counterexamples I need cases where changing the basis affects the IEM of the putatively absolutely immune self-ascription. Now I turn to discuss such counterexamples.

4. The First Counterexample to Absolute Immunity:

The Mirror Pain Case

Suppose I am usually good at knowing that I am in pain; my experience of pain reliably correlates with my pain. Still, it might not be easy to decide that one has an itch or pain for some cases. I can recognise the look of being in pain in other people and I can see that they are in pain. So when I cannot decide whether I am in pain or just have an itch I might turn to the mirror and see whether I have the

look of someone in pain. But in that case my self-ascription will not be solely based on knowing from the inside: on feeling it. It will be based on vision as well. I can have evidence for thinking that what I feel is pain if I look in the mirror like someone who is in pain. Or else, I can ask someone whether I look like someone who is in pain and thereby come to know that what I feel is pain. Call this **the Mirror Pain Case**.

I can misidentify who is in pain whenever my judgment is based on vision. In such cases what I feel is not sufficient to decide whether I am in pain. I might ask my friend whether my C-fibres are firing (if pain is C-fibre firing) and this would be another way to come to know that I am in pain and not in another state. Thus it seems I can self-ascribe pain based on non-private ways of knowing like vision or testimony.

In such circumstances I can misidentify who is in pain. If 'I am in pain' were absolutely immune this would be excluded. 'I am in pain' is one of the classical examples of an absolutely immune judgment. Consequently, if I am right that my being in pain can be known by me both from the inside and from the outside, then we have to give up the possibility of absolute immunity for this predicate.

To object that I cannot come to know that I am in pain in the same ways that other people might come to know that I am in pain would be unwise. And it seems that nobody holds this. Even Shoemaker allows that knowing from the outside can put me in a position to know that I am in pain, but I have to know this from the inside in order to self-ascribe it.³⁰

Unless we think that nothing is objective about pain we have to admit the possibility of knowing from the outside. If there is nothing objective about pain then I worry that there is nothing to be mistaken about or to know about it. In that case, pain cannot be an objective feature of the subject which is part of the world. Pain seems to be an objective feature and in most cases – evolutionarily speaking – a very useful one. (Think of pain avoidance behaviours which enable organisms to avoid danger, e.g. fire.)

³⁰ "[I]t is necessarily the case that if a person knows that a P*-predicate applies to him he knows that it applies to him in the "special way" appropriate to that predicate (which does not preclude that he should also be in a position to know that it applies to him in other ways, i.e., ways in which others might know that it applies to him)" (Shoemaker 1968: 565).

To sum up, it seems that we have found circumstances where a self-ascription 'I am in pain' – which is the classic example of an absolutely immune judgment – can be misidentified relative to 'I'. This poses a challenge to Shoemaker who holds that there can be no circumstance where an absolutely immune self-ascription is open to misidentification. Thus, I propose abandoning the Basis-independence thesis. But let me consider an objection to my strategy.

4.1. An Objection to our Strategy

Showing that one example of a putatively absolutely immune judgment is not actually absolutely immune does not show that we should abandon the notion of absolute immunity. We have seen that 'I am in pain' is not absolutely immune. This is the paradigmatic example of absolute immunity, so we, at least, have good reason to doubt that there is any self-ascription which is absolutely immune. But it would be advisable to consider more examples.

4.2. Some Other Counterexamples

Let me consider some other examples. Being out of balance or being afraid seem to be cases where if I can self-ascribe them, then I cannot misidentify whom they concern.

Consider the following cases. I am on a ship and I am overly self-confident. The captain is shouting orders to us to save ourselves from a terrible storm where the ship is in danger. I think I am doing fine. But the captain shouts at me through an earphone: 'I can see that you are out of balance. You have to go into the cabin or your life is in danger.' I self-ascribe being out of balance. But the captain might have seen someone else who is less fit for the task than me, while I am actually doing fine. Based on testimony, I can self-ascribe who is out of balance, but I may misidentify who is out of balance.

Imagine that I am climbing on a sharp stone ridge where nobody dares to go. I think I am doing fine. My friends take a Polaroid picture of me to prove that it is better for me not to climb it, because I am terribly afraid. I look terribly afraid indeed. So I self-ascribe being afraid based on my behaviour and appearance. But based on seeing an image on the picture I can misidentify who is afraid.

It is not very promising strategy to rule out the putatively absolutely immune self-ascriptions one by one. A generalised objection would be much more effective in arguing against the possibility of basis-independent absolute IEM. To present such an argument I will need to rely on what I call the Self-sufficiency principle. This principle is implicitly assumed by almost everybody in the literature, including Shoemaker. I turn to discuss the Self-sufficiency principle.

5. The Self-sufficiency Principle

Immunity to error through misidentification does not mean we have immunity from all mistakes. IEM is claimed to be a feature that is independent from infallibility, incorrigibility, self-intimation, and other special epistemic features, so it cannot require any of them (Shoemaker 1968: 556-558, 565 fn. 7, Wright 1998, O'Brien 2007, Peacocke 2008). The primary explanatory target should be IEM. Thus paradigm cases of IEM cannot be a mere consequence of another stronger epistemic feature, such as self-intimation or infallibility. Let me call this requirement of independence: **the Self-sufficiency principle**.³¹

IEM is taken to be a distinct and independent epistemic feature – and labelled as such. Let me examine the features which have to be distinguished from IEM. *Infallibility* means that when the subject judges p then the subject cannot be wrong about p . *Incorrigibility* requires that when the subject judges p , then nobody else can correct this thought besides the subject. This can be considered a form of *first person authority*. A state F (or the proposition 'I am F ') is *self-intimating* just in case whenever the subject is in a mental state F (e.g. being in pain) then she knows that she herself is F .³²

Consider *infallibility*. According to Shoemaker, one might be mistaken about the predication, being in pain. But self-ascription of this predicate is supposed

³¹ "Wittgenstein's point is not that these statements are totally immune to error, though he may have believed this to be true of some of them, but is rather that they are immune to a certain sort of error: they are immune to error due to a misrecognition of a person, or, as I shall put it, they are immune to error through misidentification relative to the first-person pronouns." (Shoemaker 1968: 556).

³² Self-intimation as discussed is related to self-knowledge and is not restricted to self-ascriptions. But our discussion requires us to discuss it only relative to self-ascriptions. There are weaker and stronger versions of these notions which may or may not be compatible with the feature being IEM. For further discussion see Alston 1971 and Tye 2009.

to be immune on all bases. Thus IEM is distinct from and cannot simply follow from infallibility.

IEM is different from *incorrigibility*. When I say 'I see a canary' then someone else can correct me that it is actually a bird of paradise. This is a paradigmatic example of absolute IEM for Shoemaker, so incorrigibility cannot be assumed.

I might think I feel an itch when I feel an ache, thus it seems that being in pain need not be self-intimating, but the judgment can still be immune. If whenever I am in a state I have to know that I am in this state (i.e. the state is self-intimating), this suffices to ensure that its self-ascription will be immune. However, if all paradigmatic cases of IEM require self-intimation then the primary explanandum would no longer be IEM, but would be self-intimation. Furthermore, according to Shoemaker, P*-predicates are those self-ascription of which is absolutely immune. However, Shoemaker explicitly denies that P*-predicates are self-intimating and this, as I have argued, seems to be correct. Thus paradigmatic examples of IEM cannot be the self-ascription of features which are self-intimating.

Accordingly, the Self-sufficiency principle ensures that IEM is not merely a derivative feature – a consequence of another more fundamental feature of the use of 'I'. The Self-sufficiency principle is a point of common agreement even though it has never been explicitly stated in the literature, and it will play an important role in the discussion below.³³ With this in place, I can now turn to the generalised objection against basis-independence, which undermines the possibility of absolute IEM.

6. Generalised Objection to Absolute Immunity:

What is Implausible in the Idea of Absolute Immunity?

I will argue that unless P*-predicates are self-intimating, their self-ascription cannot be immune on all possible bases. (In Shoemaker's framework, P*

³³ Dialectically note that these features cannot be required by IEM if IEM is the primary explanandum. If, in a paradigmatic case, an immune judgment (*Fi*) required that *Fi* is infallible, incorrigible or self-intimating, then the primary explanandum would **not** be IEM but these other features. IEM follows from infallibility; if a judgment is infallible it follows that it is immune. But if IEM would require infallibility then infallibility would be the primary explanandum. This is also the case with self-intimation and incorrigibility.

predicates are a subclass of psychological predicates and their self-ascription is absolutely immune.) Before presenting my argument, let me emphasise that Shoemaker endorses the Self-sufficiency principle (Shoemaker 1968: 556-558) and does not require that P*-predicates are self-intimating: “[S]elf-ascriptions of P*-predicates need not be incorrigible and (...) it is not necessarily the case that if a P*-predicate applies to a person that person knows that it applies to him [i.e. P*-predicates need not be self-intimating].” (Shoemaker 1968: 565).

The general problem that makes absolute immunity unattainable is as follows. On the supposition that absolute immunity is Self-sufficient (that it neither requires infallibility nor self-intimation) we cannot exclude the possibility that I can self-ascribe *F*-ness (for all predicates) based on a non-private way of knowing. Based on a private way of knowing like introspection, my self-ascription is immune. Based on a non-private way of knowing (e.g. vision) the self-ascription is not immune. But basing the self-ascription on a non-private way of knowing cannot be excluded unless we require that *F*-ness is self-intimating. This, however, is ruled out by the Self-sufficiency principle. For example, *being in pain* cannot be self-intimating if its self-ascription is a paradigmatic example of absolute immunity according to the Self-sufficiency principle.

Take any paradigmatic case of absolute IEM, *Fi* (B). Properties of mine can be self-attributed from the inside or from the outside according to Shoemaker. It is not only other people who can attribute that I am *F* from the outside; I too can believe that I am *F* from evidence from the outside. Only self-intimation of *F*-ness ensures that I know that I am *F* whenever I am *F*. So if *F*-ness is not self-intimating (as it has to be for a paradigmatic case of immunity) then I might not believe that I am *F* when I am *F*. But in such a case I might come to believe that I am *F* in the same way as others believe that I am *F* – from the outside. Anyhow, if others can believe ‘I am *F*’ through a non-private way of knowing (from the outside) then nothing can rule out that I can know (or at least believe) through such ways that ‘I am *F*’. And even Shoemaker allows that I can know from the outside that I am *F* (Shoemaker 1968: 565).

Let me run through this reasoning for the judgement ‘I am in pain’, which is the paradigm case of an absolutely immune judgement:

1. I can think that I am in pain based on feeling this from the inside.
2. Others can think that I am in pain based on vision (from the outside).
3. If others can think that I am in pain based on vision, then I can also believe this based on vision (e.g. looking at my reflection or image)
4. If I think that I am in pain based on vision then I can be mistaken about who is in pain.
5. When I do not know whether what I feel is pain I can come to think that I am in pain based on vision.

But then I can misidentify who is in pain based on vision in such cases.

Often we are not sure what state we are in and we need further evidence to decide. I might feel something but I do not know whether I am in pain as opposed to some other kind of discomfort. One could think that self-ascription of pain is necessarily from the inside. But when I think based both on private and non-private ways of knowing that I am in pain then I can misidentify who is in pain. At least when my judgement is based both from inside and outside then it can turn out that what I feel is not pain but the one whom I look at is in pain. In this case I can misidentify who is in pain.

But, if I do not know that I am *F* when I am *F* (*F*-ness is not self-intimating), then I might come to believe from the outside and the inside. Thus, for a putatively absolutely immune self-ascription, which is not self-intimating, I might come to believe that I am *F* from the outside and the inside. If I can believe that I am in pain based on evidence from the outside, then I might misidentify who is in pain. I might come to believe that I am *F* from the outside when someone else is *F*, but I am not (feeling something else). But then I can misidentify who is *F*, when I come to believe it from the outside and from the inside. The counter-examples to absolute IEM demonstrate this possibility.

Consequently, either there is no absolute immunity or we have to reject the Self-sufficiency principle. Rejecting the Self-sufficiency principle is not yet sufficient for defending absolute immunity, because one would also have to argue that *P**-predicates are self-intimating. I do not see enough reason to accept this; my pain might be in the background of my attention and I only

come to realise it later. For example, I might remember later that I already had a headache when I was so angry with someone – but I did not realise it then. This is possible even if pain requires pain phenomenology. One may argue that it is only confabulation for unjustified anger. However, I might not pay sufficient attention to my headache. I might have a headache but I am not fully aware of it because I am, for example, trying to finish my argument. I can redirect my attentional resources to something else than the headache.

7. The Fall of Basis-independence

According to the basis-independence thesis, an absolutely immune self-ascription is immune in all circumstances, so it should be immune on all possible bases. Basis-independence follows from circumstance-independence. Thus to deny the basis-independence thesis is to deny the circumstance-independence thesis and thereby to deny absolute immunity.

According to the basis-independence thesis, a supposedly absolutely immune self-ascription, *Fi*, cannot be misidentified on any basis. So if we find that there are always such bases on which *who is F* can be misidentified, then we should deny the basis-independence thesis and thereby deny absolute immunity. This is what we have seen. I have shown that the self-ascription might be based only on non-private, or jointly on private and non-private ways of knowing – unless we presuppose that all *P**-predicates are self-intimating. But we have no reason to assume this. If *P**-predicates should be self-intimating, this would violate the Self-sufficiency principle. Thus we have no reason to believe in absolute IEM, unless we are prepared to give up the Self-sufficiency principle. But to give up Self-sufficiency is to give up IEM as we have conceived it.

Let me summarise my main argument against absolute IEM. Suppose *P*i* is a paradigmatic case of absolute IEM. Nothing excludes that I know that ‘I am *P**’ in the same way as others may know it – from the outside, because (i) for any predicate being instantiated by me, it can, in principle, be known from the outside and (ii) such predicates (*P**) cannot be restricted to those where being *P** is self-intimating (by the Self-sufficiency principle). Whenever *Fi* is based on knowing from the outside, the subject can misidentify *who is F*. I might not

know that I am *F*. This cannot be excluded unless we give up the Self-sufficiency principle and provide reason to believe that *P**-predicates are self-intimating or their self-ascription is infallible. But we have no reason to give up Self-sufficiency and believe *P**-predicates are self-intimating or their self-ascription is infallible. This provides compelling reason to believe in the basis-relativity of all kinds of immunity.

8. The Special Basis Explanation

Let me take a step back and try to reconstruct Shoemaker's explanation of absolute IEM. Surprisingly, this will enable me to find an explanation of the IEM of self-ascriptions which accommodates basis-relativity. For the sake of reconstructing Shoemaker's explanation of absolute IEM, let us forget that absolute IEM is not possible.

How does Shoemaker try to explain absolute IEM? If absolutely immune *Fi* is immune on all bases, then how might the possibility of misidentification be ruled out for *Fi*? To think that a self-ascription is, on all possible bases, immune to error through misidentification requires an explanation of how this is possible. Shoemaker's thought is that this is possible because the basis is special. Such kinds of bases should be such that they exclude the possibility of misidentification. Furthermore, it requires an explanation which excludes the possibility that the self-ascription can be made on a basis which allows misidentification. So there should be a special basis which excludes the possibility of misidentification and the self-ascription should be made only on such a basis. Call this line of thought that Shoemaker seems to be relying on **the maximal special basis explanation**. As we will see this explanation has a grain of truth in it and guides us in the good direction.

What makes it possible that the self-ascription of some predicates (*P**-predicates) is – apparently – absolutely immune? Absolute immunity requires that such self-ascriptions are made on a special basis (Shoemaker 1968: 565). The basis is, minimally, the relevant way of knowing on which the judgment is based, like proprioception, introspection, vision, and testimony. What makes the basis special?

Shoemaker is continuing the discussion begun by Wittgenstein and Anscombe of how the subject might come to know of the material object which she is. He points out that there are special bases through which the subject can know **only** of herself (or at least introspection is such). If, for the subject, there are private ways of knowing of the subject (which cannot be a way of knowing of anything but the subject) and self-ascription of P*-predicates requires that the subject can self-ascribe a P*-predicate *only* based on a private way of knowing, then this can explain how absolute immunity is possible. Shoemaker expresses this idea in the following passages:

And even if my "self" is a flesh-and-blood person, why shouldn't it be accessible to me (itself) in a way in which it is not accessible to others, so that in knowing that what is presented to me is presented in this special way – from the inside, as it were – I would know that it can be nothing other than myself? (Shoemaker 1968: 562).³⁴

It has often been held to be one of the defining features of the realm of the mental, or the psychological, that each person knows of his own mental or psychological states in a way in which no other person could know of them. We can put what is true in this by saying that there is an important and central class of psychological predicates, let us call them "P*-predicates", each of which can be known to be instantiated in such a way that knowing it to be instantiated in that way is equivalent to knowing it to be instantiated in oneself. (Shoemaker 1968: 565, my underlining).

For Shoemaker, the self-ascription of P*-predicates is absolutely immune. P*-predicates cannot be self-ascribed unless they are self-ascribed based on a private way of knowing. This explains the IEM of self-ascriptions:

³⁴ "Now there is a perfectly good sense in which my self is accessible to me in a way in which it is not to others. There are predicates which I apply to others, and which others apply to me, on the basis of observations of behaviour, but which I do not ascribe to myself on this basis, and these predicates are precisely those the self-ascription of which is immune to error through misidentification." (Shoemaker 1968: 562, my underlining).

Maximal special basis explanation: *Self-ascription of F is absolutely immune iff F cannot be self-ascribed unless it is self-ascribed based on a private way of knowing.*

A special way of knowing is private. Private ways of knowing are such that only the subject can employ them to know of the subject. Proprioception provides the sense of the position of one's limbs relative to one's other limbs and other parts of one's own body.³⁵ That my legs are crossed is something that might be known solely based on proprioception. Proprioception is private; nobody except me can know of that object which is me through it. Kinaesthetic sense provides awareness of the movement of the subject's body.³⁶ Only the subject can employ it to gain knowledge of the subject. Accordingly, examples of private ways of knowing are proprioception, kinaesthesia, introspection and channels for pain/balance/heat. In contrast, vision, audition, testimony and inference are non-private ways of knowing, because other people can also employ them for knowing about the object which is me.³⁷

The suggestion is that self-ascriptions based only on a private way of knowing will be immune. According to Shoemaker, absolute IEM occurs because some predicates can only be known for the subject through a private way of knowing. On this picture, private ways of knowing and their features explain why certain self-ascriptions are absolutely immune. Interestingly, giving up the possibility of absolute IEM leaves a weaker version of the special basis explanation untouched.

³⁵ A private way of knowing need not be a non-perceptual way of knowing. Let me allow, contrary to Shoemaker, that proprioception and channels of pain are not just examples of private ways of knowing, but are genuine perceptual ways of knowing.

³⁶ Evans criticized Shoemaker for assuming that only self-ascription of some psychological predicates (P*-predicates) are immune, but there are no immune bodily self-ascriptions. Evans correctly observed that based (only) on proprioception self-ascriptions are immune. The special basis explanation allows Shoemaker to answer Evans. Such bodily ways of knowing (proprioception, kinaesthesia) are special private ways of knowing; thus based only on them self-ascriptions are immune.

³⁷ I take 'object' as a primitive for things which have a certain unity and boundary, endures in time, possess 'certain features, and may have parts.

8.1. Knowing from the Inside and Private Way of Knowing

I want to suggest that Shoemaker is correct in some aspects of his Special Basis explanation, but wrong about denying basis-relativity. A weaker version of the Special Basis explanation does not exclude basis-relativity. According to this, a self-ascription (B) is immune (in a circumstance) if it is only based on private ways of knowing (in this circumstance). But it is not immune if it is based on a non-private way of knowing. I simply call this minimal special basis explanation: the **special basis explanation**.

The special basis explanation: *The self-ascription of F (B) is immune iff B is a private way of knowing.*

According to the special basis explanation, when the subject only knows/believes from the inside *Fi* (B), then the self-ascription on this basis, B, is immune.³⁸ We expect 'I have my legs crossed' will be immune based solely on proprioception, but based on vision it will not be immune. This is exactly what we find. When a subject knows something through a private way of knowing then the subject knows it from the inside. At this point, I only find this claim very plausible. I will develop it further in chapter 8.

9. Summary

IEM is agreed to be Self-sufficient; paradigmatic cases of it cannot require infallibility or self-intimation. I called this the Self-sufficiency principle. According to Shoemaker, absolutely immune self-ascription is immune in all circumstances. Absolute IEM of *Fi* requires that *Fi* in all circumstances, accordingly on all bases, has to be immune relative to 'I'. I argued that unless *Fi* is self-intimating (or infallible) this could not be secured. But being self-intimating (or infallible) is excluded by the Self-sufficiency principle. Either one has to give up the Self-sufficiency principle or the possibility of absolute IEM. The Self-sufficiency principle is the only way to secure that IEM is an independent and not merely a derivative feature of uses of 'I' for which it was

³⁸ There are potential counterexamples: 'I see a canary', based on vision, is supposed to be immune and 'I face the table', based on vision, could be immune. See chapter 10 (*Proffitt-Shoemaker objection*) for my response.

introduced. Moreover self-intimation (or infallibility) of immune self-ascriptions has not been established and does not seem likely. This provides a general argument for the basis-relativity of the immune of self-ascriptions.

Toward the end of the chapter, I discussed Shoemaker's special basis explanation of IEM. I suggested that Shoemaker is correct about some aspects of his Special Basis explanation, but wrong to deny basis-relativity. There is a weaker version of his special basis explanation which I find powerful. According to this weak version, the self-ascription of F (B) is immune if B is a private way of knowing.

Chapter 5

Is immunity to error through misidentification immune to empirical counterexamples?

Abstract

In this chapter, I will consider three putative counterexamples to immunity to error through misidentification and argue that they fail. The counterexamples are intended against the phenomenon described by Shoemaker (1968), but I argue that they only work against a particular explanation of IEM of self-ascriptions.

The first putative empirical counterexample is due to Campbell (1999b) and the second and the third are due to Lane and Liang (2011). I will show that none of the empirical counterexamples work, though they fail for different reasons. Obviously this does not mean that there can be no counterexamples to immunity. The aim of this chapter is not just to argue against the putative counterexamples to IEM but also to isolate the real issue behind them. In answering these challenges we will achieve a better understanding of the phenomenon of IEM.

1. The first putative counterexample against immunity

John Campbell argued that inserted thought is a counterexample against the phenomenon of IEM in his paper 'Schizophrenia, the Space of Reasons and Thinking as a Motor Process' (1999b). Schizophrenic patients suffering from inserted thoughts think that someone inserts thoughts into their head. Campbell argues that inserted thought is a counterexample to IEM:

What is so striking about the phenomenon of thought insertion as described by schizophrenic patients is that it seems to involve an error of identification. The patient might say, 'Thoughts come into my head like

“Kill God”. It’s just like my mind working, but it isn’t. They come from this chap, Chris. They’re his thoughts.’ (Frith 1992, 66). A patient who supposes that thoughts have been inserted into his mind by someone else is right about which thoughts they are, but wrong about whose thoughts they are. (Campbell 1999b: 609).

Campbell claims that the difficulty for IEM arises from the fact that patients with inserted thought are sometimes wrong about who is the agent, the thinker of their thought. They think the thinker of their thought is someone else. Campbell’s conclusion is that we have to recognise that immunity has enabling conditions. I will argue that Campbell is mistaken: it is the capacity to self-ascribe based on knowing from the inside that has enabling conditions, and not the IEM of self-ascriptions.

There are four problems with Campbell’s argument. The first is that the argument which is supposed to be against Shoemaker is not even spelled out. The second is that we need a putative counterexample *against the immunity of a self-ascription*; in Campbell’s paper the relevant self-ascription is missing. ‘Kill God’, his example, cannot play that role. It is not a self-ascription. In a way, the first and the second problems are superficial. I will answer them on behalf of Campbell. But there are two other worries, and they will disqualify the putative counterexample against IEM.³⁹ Let me proceed by reconstructing Campbell’s argument.

1.1. The Reconstruction of Campbell’s argument

Shoemaker claims that, for an immune self-ascription, ‘I am in pain’ based on introspection,⁴⁰ it cannot be the case that when it turns out that I am not in pain, then, on the same basis, I can think that someone else is in pain.

³⁹ Gallagher (2000) and Coliva (2000) have responded to Campbell on the basis of assuming a particular explanation of IEM, the subject use of ‘I’, which, on my view, is about the ownership but not about the agency of thought. However, Campbell is discussing agency and not ownership. My response does not assume any particular interpretation or explanation of IEM. I show that there is a structure necessary for misidentification which is lacking in Campbell’s counterexample.

⁴⁰ The self-ascription of pain seems to be based on perception and not on introspection. However, this does not matter for my discussion in this chapter. For the sake of argument let me stay with Shoemaker’s original idea.

For a counterexample against the immunity of self-ascriptions, we need at least to have a self-ascription. However, this is missing in Campbell's paper. But it is a prerequisite for a counterexample against the IEM of self-ascription that one provides a relevant self-ascription, with its basis, which will turn out not to be immune on that basis. The best candidate for Campbell's required self-ascription is 'I think that p '. This can be read as my believing that p or that I am the thinker of p (Crane 2011). I only mean the latter: being the thinker of the thought/agent of the thinking.

What is the basis of the self-ascription? Campbell's answer is this: "the schizophrenic has introspective knowledge of a thought of which he does not recognise himself to be the agent" (Campbell 1999b; my emphasis). The basis is introspection. But what can be misidentified? Campbell's idea is perhaps that an inserted-thought patient can misidentify *who thinks that p* based on introspection. This is because an inserted-thought patient sometimes thinks that he is not the thinker of his thought. Accordingly, 'I think that p ' is not immune relative to 'I' in the case of an inserted-thought patient. Thus it seems that we have a counterexample to the IEM of 'I think that p ' based on introspection.

1.2. Criticism of Campbell

Suppose that an inserted-thought patient says 'I think that marmalade is bad' – based on introspection, and the patient continues: 'But Chris inserted this thought into my head'.⁴¹ Thus such a patient misidentifies *who thinks the thought marmalade is bad*. For an immune self-ascription, like 'I am in pain' based on introspection it cannot be the case that when it turns out that I am not in pain, then, on the same basis, I can think that someone else is in pain. The structure of a counterexample has to look like this: there is a self-ascription made on a basis B , but it involves a misidentification because the subject believes that b (which is F and which is in fact a distinct individual) is she herself. Immunity excludes the possibility that a self-ascription based on basis B

⁴¹ It is not important whether something was ever self-attributed in exactly this manner by an inserted-thought patient. The problem is that such a self-attribution and a denial is possible for an inserted-thought patient while IEM is supposed to exclude this possibility, according to Campbell.

involves a misidentification of ‘who is *F*’. So for a counterexample to a claim that a self-ascription is not immune, the self-ascription has to be based on an identification of ‘who is *F*’ in order to demonstrate that misidentification is possible.

In the case of misidentification, I judge *Fa* (B) on the basis B which is for judging *Fb*. In chapter 3, when I explicated the basic notion of misidentification, I called *Fa* (B) ILL, because thinking *Fa* involves the subject presupposing that *a* refers to *b* (which is *F*) when this is not the case. I called *Fb* WELL, because it is of *b* and there is no mistake of identification involved in this. (It is allowed that *b* is not *F*.) As we saw in chapter 3, the structure of the misidentification is this: I have an ill-grounded thought involving a misidentification *Fa* (B) when I am in a position to form a well-grounded thought about *b*, *Fb* (B).

Now let us see whether this structure of misidentification is present in Campbell’s putative counterexample. In our reconstruction of Campbell’s case the self-ascription ‘I think that marmalade is bad’ is well-grounded; it correctly concerns me. As I have noted, for a counterexample to a claim that a self-ascription is not immune – i.e. in order to have a case of misidentification – the self-ascription has to be based on a misidentification of who is *F*. The revision, ‘Chris thinks that marmalade is bad’, is the one which is ill-grounded and involves a misidentification (if any judgement here involves misidentification at all). But in that case Campbell’s case has the wrong structure.

The self-ascription should involve a misidentification. But the self-ascription in Campbell’s case does not involve misidentification. If anything, the judgement ‘Chris thinks that marmalade is bad’ involves a misidentification. But IEM excludes the possibility that the self-ascription on a particular basis could involve a misidentification. The IEM of self-ascription is about a self-ascription excluding a particular kind of mistake – misidentification – in the self-ascription. IEM is not about ascriptions to others, like ascriptions to Chris, but only about self-ascriptions. To sum up, a counterexample would require the opposite structure. The self-ascription should involve a misidentification on a particular basis B (ILL) and the revised judgment should not (WELL).⁴²

⁴² Evans (1982) and Pryor’s *de re* IEM have similar structure: there is an ill-grounded self-ascription and a well-grounded thought of something other than the subject.

The counterexample doesn't work even if we analyse it in terms of the dominant account from Evans in terms of identification-freedom (see chapter 6). For Evans, misidentification requires thinking Fa because the subject presupposes an identification component with the logical form $a=b$ (Evans 1982: 180-181). An 'a'-thought (Fa) will be immune relative to a if the knowledge of Fa does not depend on the truth of an identification component (with the logical form: $a=b$). In contrast, misidentification happens where the identification component is false (a is not identical with b). On Evans's characterisation the inserted-thought patient should think that he is Chris if 'Chris thinks that marmalade is bad' is based on a misidentification. But this is not the case. Thus no thought involves a misidentification *a la* Evans. But Campbell's target is Shoemaker, so I can put aside this issue.

1.3. The main criticism

The main difficulty is that in the counterexample the self-ascription is not based on a misidentification – as it is required for a counterexample to the IEM of self-ascriptions. Immunity excludes the possibility that on the relevant basis it might not be me, but someone else who is F (has messy hair). I judge Fa (B) when B is for judging Fb . Thus for a counterexample to an immune self-ascription the relevant basis has to be such that I can know of another object and its F -ness (messy hair) through it. Shoemaker puts this point as follows: "... this is perhaps the most important point, the identification of a presented object as oneself would have to go together with the possibility of misidentification, and it is precisely the absence of this possibility that characterizes the use of 'I' that concerns us" (Shoemaker 1968: 561). Thus according to Shoemaker, for the possibility of misidentification, there should be an object other than myself which is present to me, and I have to believe that it is me. No other object but myself can be presented to me through introspection. Accordingly, the structure of a counterexample should be this: there is a self-ascription made on a basis B, and this involves a misidentification because the subject believes that b (which is F) is she herself. But this cannot happen if the judgment is based only on introspection according to Shoemaker.

Is it the case that for misidentification, the other object (which is not me, but is F), b , has to be present to the subject? Yes! According to Shoemaker: for Fa , based on a special way of knowing, W , there cannot be an object other than myself, b , which is present to me through the same private way of knowing, W , through which *I am presented to myself* (Shoemaker 1968: 565). Thus if Fa is based only on a special private way of knowing then I cannot misidentify which object is F . To make it simple, a self-ascription is immune iff a way of knowing W is such that if the subject knows through W that F is instantiated then the subject knows that she herself is F . This cannot be the case unless nothing but one object can be known through W . As Shoemaker writes: "We can put what is true in this by saying that there is an important and central class of psychological predicates, let us call them "P*-predicates," [a subset of psychological predicates the self-ascription of which is or might be immune] each of which can be known to be instantiated in such a way that knowing it to be instantiated in that way is equivalent to knowing it to be instantiated in oneself" (Shoemaker 1968: 565). The subject knows that it is of herself because knowing that F is instantiated through such special ways of knowing involves knowing that it is instantiated in herself. Thus no other object can be known through such a way of knowing, otherwise it would **not** be excluded that one knows of another object through W .

In sum, what we need for misidentification in a self-ascription is the following. An object mistakenly has to be presented under the first person mode of presentation, as me, even though she is not me. *Chris* in the example is not presented to the schizophrenic from the first person perspective "as me" in 'I think that marmalade is bad'. Therefore, there is no misidentification on which the self-ascription is based. The IEM of self-ascriptions (B) concerns only self-ascriptions (B). Whatever mistake we have been considering here happens in the other-ascription (an ascription to someone other than myself). However, the IEM of self-ascription does not predict anything about other-ascription of the relevant property.

1.4. Why does inserted thought matter if it is not a counterexample to the immunity of certain self-ascriptions?

Still, there is a grain of truth in the putative counterexample. I want to suggest that instead of being a counterexample against immunity, Campbell's example might work against a *particular explanation* of immunity. The counterexample might be directed against the thought that known through introspection, it is necessary that the subject self-ascribes what is known through it (Shoemaker 1968: 565). This is a possible explanation of IEM.

Call internal ways of knowing (though which the subject knows only of the subject herself – like proprioception or introspection) **ways of knowing from the inside**. When I know from the inside, then I cannot be wrong about *who is F*. Against this explanation of IEM, 'Chris thinks that Marmalade is bad' (B=introspection), is a counterexample. It is a counterexample against the claim that knowing something from the inside ensures that one knows it of oneself. It is not sufficient to know from the inside in order to self-ascribe it. There should be a mechanism which enables me to self-ascribe it. This mechanism – an enabling condition of the integrity of the capacity for self-ascription – can malfunction.⁴³ This suggestion of Campbell's is correct and insightful, but it is not about IEM. Rather it is about whether one necessarily self-ascribes whatever one knows from the inside. This is not IEM. It could only be an explanation of it. Thus it is only a counterexample against **an explanation of immunity** and not against the phenomenon of the immunity of certain self-ascriptions.

Shoemaker seems to claim that when a self-ascription is based on introspection then one necessarily knows who is *F* and necessarily self-ascribes *F* (Shoemaker 1968: 563-564). This requires that being *F* known through introspection is self-intimating; whenever the subject knows of *F*-ness through introspection then she knows that she herself is *F*. But then we have a counterexample against this explanation of immunity of self-ascriptions but not against the phenomenon itself. (If there were no other possible explanation of immunity except this one then perhaps we would accept inserted thought as a

⁴³ See McDowell 1994 and Burge 2010 for discussions of the contrast between enabling and constitutive conditions.

counterexample to IEM. However, there are other explanations.) Campbell is correct that there are enabling conditions for self-ascription, and, for the proper functioning of self-ascription, the mechanism enabling it has to work properly.

What makes Campbell's worry really interesting is that inserted thought is a counterexample against the claim that one cannot be wrong about who is the thinker of his own thought. All the same, this is an independent question from the IEM of certain self-ascriptions. The complexities here have to be dealt with elsewhere.⁴⁴

The main issue behind Campbell's worry is that there is a mechanism that enables self-ascription of properties known from the inside. This can malfunction. The subject can know something from the inside without self-ascribing it. Thus it is the capacity to self-ascribe that has enabling conditions and not the IEM of self-ascriptions. This is a lesson to take on board even if inserted thought is not a counterexample against the immunity of 'I think that p' based on introspection.

2. The second putative counterexample to immunity

Lane and Liang (2011) provide the remaining two putative empirical counterexamples to the phenomenon of immunity. They claim that: "[a]lthough some aspects of Shoemaker's views on immunity have been disputed, IEM itself has never been severely threatened by any empirical challenge" (Lane and Liang 2011: 78). Thus the target is the phenomenon of IEM and not an explanation of IEM. As I will argue, IEM is not yet challenged by the empirical counterexamples. I turn to discuss the two putative empirical counterexamples.

The first putative counterexample against IEM consists of judgments made by a somatoparaphrenic patient FB. Patients with somatoparaphrenia think that one of their hands (usually the left hand) belongs to someone else. FB thinks that her left hand is her niece's hand even when she sees it. In an experiment (Bottini et al. 2002) the patient FB always claimed that her niece was touched when she was touched on her affected, somatoparaphrenic hand based on

⁴⁴ Roessler's essay 'Thought insertion, self-awareness, and rationality' (2014) tries to find what makes thought insertion so surprising. We seem to presuppose that S cannot think unless S knows under the first person mode of presentation that S is thinking, at least implicitly. My only disagreement with his claims in this essay is that thought insertion is a counterexample to IEM.

somatosensory stimulation. Bottini and colleagues observed that FB was never able to report that *she* was touched on her left, affected hand based on somatosensory stimulation (Bottini et al. 2002: 251). However, they also conducted trials on which they warned FB that her niece's hand would be touched and after that they asked her whether her niece was touched. Throughout all the testing conditions, FB was blindfolded, so she could not see but only feel her hand. On the occasions when FB was forewarned that her niece's hand would be touched, FB always reported correctly that the hand which she claimed to belong to her niece (but was actually FB's hand) was touched. As a control, catch trials were introduced where the experimenter claimed that she would touch FB's niece, but there was no tactile stimulation whatsoever. In such cases FB reported that her niece was not touched. Thus, as I understand it, FB knows *de re* of a particular thing that is being touched, but under the wrong mode of presentation: 'my niece'. She *de re* knows of herself but under the wrong mode of presentation – not as herself.

According to Lane and Liang this is a counterexample against immunity or, at least, Shoemaker's version of immunity. The reason is that "... FB commits an error through misidentification regarding just who is the subject of the sensation" (Lane and Liang 2011: 86). It is not yet clear how 'my niece was touched' is a counterexample to immunity. Lane and Liang (2011: 87) emphasise that Shoemaker's holds that: "in being aware that one feels pain one is, tautologically, aware, not simply that the attribute *feel(s) pain* is instantiated, but that it is instantiated *in oneself*" (Shoemaker 1968: 563-564).

It is true that Shoemaker holds that knowing from the inside necessarily means knowing of myself. I discussed this above. However, this is only an explanation of IEM. This explanation could be revised to deal with worries like that. But more importantly these quotes do not support Lane and Liang's claim that being the subject of an experience explains IEM.

2.1. Objections against Lane and Liang

Lane and Liang hold a specific explanation of IEM and they assume that Shoemaker shares their view. There are three points against them. First, there are several irresolvable difficulties in thinking that being the subject of the

experience explains immunity. I am always the subject of my experience, my perceiving, my imagining, and my thinking, but not all self-ascriptions based on these are immune. If I look at the mirror and, based on this perception, I judge that my legs are crossed, the judgment will not be immune – but I remain the subject of my perception.

They could suggest that self-ascription of the property of being the subject is immune. However, I need not self-ascribe being the subject of experience in order to get IEM. The judgement ‘I have my legs crossed’ based on proprioception is immune but based on vision is not IEM. The difference cannot be accounted for by saying that I am the subject of experience in one case, but not in the other case. Similarly, it cannot be that in one case I self-ascribe being the subject of experience but not in the other case.⁴⁵

Second, they do not attempt to show that Shoemaker shares this explanation of IEM.

Third, there is no self-ascription which is supposed to be immune but as they argue, turns out not to be immune. The putative counterexample is supposed to be against IEM and not against an explanation of IEM (Lane and Liang 2011: 78). The difficulty is that immunity is a feature of self-ascriptions *Fi* (B), so the counterexample to IEM should have a self-ascription and a basis. When one argues against the immunity of a self-ascription then one first needs a self-ascription and its basis on which *it is supposed to be* immune, but which on this basis turns out to be not immune.

Thus we need a self-ascription (B) in which who is *F* can be misidentified but the self-ascription is supposed to be immune. However, ‘My niece’s hand was touched’ is not a self-ascription. Moreover, FB is completely unable to make self-ascriptions about herself being touched when her left hand is being touched. FB was never able to report herself being touched based on somatosensory stimulation of the affected somatoparaphrenic hand (Bottini et al. 2002: 251). Thus we do not have a counterexample against the immunity of a self-ascription (B). There is no self-ascription which is supposed to be immune on a basis, but turns out not to be immune. Consequently, we have no counterexample to IEM.

⁴⁵ Liang is working on this explanation in his forthcoming paper.

2.2. Suggestion to Lane and Liang

Lane and Liang offer putative empirical counterexamples to the phenomenon of immunity. Instead, I suggest the counterexample should be offered against an explanation of IEM. Shoemaker's explanation of IEM would be a better target to criticize than the phenomenon of IEM. For Shoemaker, knowing from the inside necessarily means knowing of myself.⁴⁶ Shoemaker – as we have seen – holds the Special-basis explanation, and this explanation would be a much better target for them to attack. FB knows through internal ways of knowing of the object which is her (*de re*) that it is touched – but she ascribes being touched to her niece. So let me stick to what Shoemaker holds, to see what can be criticised on the basis of the discussed empirical experiment.

When one argues against the immunity of a self-ascription then one first needs a self-ascription and its basis on which *it is supposed to be* immune, but which on this basis turns out not to be immune. This will be a challenge for all explanations of IEM since it is a counterexample to the phenomena itself. In contrast, when one argues against an explanation of IEM then the strategy is different. One needs to provide a counterexample to the explanation. But we are supposed to have a counterexample against the phenomena and not against an explanation of IEM. This cannot be done unless we have a self-ascription which is supposed to be immune but, as they argue, turns out not to be immune.

The mistake might be that one firmly believes in an explanation of immunity and disbelieves other possible explanations. But if one argues against an explanation of immunity the explanation could be simply revised. To argue against immunity we need a counterexample which is independent from whatever explanation of immunity one adores.

Lane and Liang could have been taking themselves to be arguing against an explanation of IEM and not against the phenomenon of IEM. In this case my objection would not apply. But they make it clear that this is not the case: "Although some aspects of Shoemaker's views on immunity have been disputed, IEM itself has never been severely threatened by any empirical

⁴⁶ For the sake of argument I will ignore the fact that Shoemaker requires that the subject has to know in which way (e.g. introspection/from the inside) she knows of *F*-ness.

challenge” (Lane and Liang 2011: 78). For this they need a self-ascription (B) as a counterexample.

To sum up, to find a counterexample of immune self-ascription, we at least need a self-ascription. To have a counterexample against an explanation (which can be revised or abandoned) of IEM of self-ascriptions is different from having a counterexample against IEM of self-ascriptions.

3. The third putative counterexample to immunity

The third putative counterexample against immunity is a judgment made after experiencing the body swap illusion. Let me describe this illusion first. The body swap illusion (Petkova and Ehrsson 2008) is based on using virtual reality techniques in exchanging the point of view of the experimenter and the participant. A video camera on the head of the experimenter records what is seen from his perspective and a head mounted display (HMD) presents the participant with what the camera projects from the head of the experimenter. The hand which the participant sees through his HMD is anatomically congruent with his own hand, a condition which is known to be a requirement on such illusions (Tsakiris 2010), but this is actually the experimenter’s hand. The participant sees the experimenter’s body from a first-person, egocentric point of view as when one looks down on his or her own body. During the course of experiment the participant and the experimenter either squeeze one another’s hand synchronously or, in the control condition, asynchronously. The illusion – feeling the body seen from the egocentric point of view to be the subject’s own body – appears in the synchronous condition.

Some participants spontaneously mentioned after the experiment: ‘I was shaking hands with myself.’ The judgment ‘I was shaking hands with myself’ was reported after the experiment not as part of the experiment and merely an anecdotal remark; as such, it might be metaphorical or a joke. Speaking about ‘shaking hands’ is already pretty metaphorical. It is actually just squeezing the hand. But for the sake of argument let us take the report at face value and see if it can work.

The putative counterexample is the following ‘I was shaking hands with myself’, with vision as its basis. Finally, we have a case that satisfies the basic

formal requirement for a counterexample: there is at least a self-ascription (B).

According to Lane and Liang there is a difficulty for IEM: “This poses a problem for Shoemaker’s IEM. From the subjects’ first person point of view, the handshaking experience involves the awareness that I am the agent of this action. This recently has been called “agentive experience” or “agentive self-awareness” – I experience myself as someone who is doing something. What is distinctive about the Body Swap Illusion is that the subjects’ agentive experience is mistaken. Although it was really the experimenter who was shaking their hands, the subjects misrepresented themselves as the agent of the action.” (Lane and Liang 2011: 92).⁴⁷

The most important objection to Lane and Liang is that if they are right, then according to Shoemaker ‘I was shaking hands with myself’ based on vision should be immune. For an immune judgment ‘I am *F*’, based on B, it cannot be misidentified *who is F*. This means that Shoemaker has to hold that the subject cannot misidentify who was shaking hands with him based on vision. But Shoemaker holds the opposite view.

What is Shoemaker’s view on misidentification based on vision? Recall ‘I have messy hair’ based on vision is not IEM for Shoemaker. Wittgenstein has an example where judging ‘I am bleeding’ or ‘I have my arm broken’ based on vision of a hand is not IEM because it might be someone else’s hand (Wittgenstein 1958: 66-67). Shoemaker repeats Wittgenstein’s example. So neither Wittgenstein nor Shoemaker thinks that based on vision I cannot misidentify myself. These are the paradigmatic cases when I *can* misidentify myself.

Let us try to amend Lane and Liang’s example against this worry. What is the basis and what is the self-ascription? “When they experience the illusion of shaking hands with themselves, their experiences involve misrepresentation of action awareness – that is, the misrepresentation concerns “who” shakes their hands. This poses a problem for Shoemaker’s IEM.” (Lane and Liang 2011: 92).

⁴⁷ The body-swap illusion seems to work against an example of an immune self-ascription from Peacocke (2012: 145). According to him, ‘My arm is broken’ is immune based on “visual experience of [an] arm, seen as attached to one’s own body”. In the body-swap illusion setup, I see the experimenter’s arm as attached to my own body. Imagine seeing a broken arm in this setup and forgetting that one is in a virtual reality setup.

So they might think ‘I was shaking hand with myself’ means I did it based on agent awareness, and this is supposed to be immune relative to ‘I’. (Let us bracket that the judgment is based on vision as well.⁴⁸) But indeed the participant did it. More importantly the participant squeezed the experimenter’s hand and supposedly has agent awareness for it. Recall, the body swap illusion was only present in the synchronous condition. In the asynchronous condition they squeeze at different times and there was no body-swap illusion present. This means that there was no case when the experimenter performed an action, but where the participant did not perform a similar action and the body-swap illusion was experienced. The asynchronous squeezing condition is the control condition for showing that the body-swap illusion only comes with the synchronous condition.⁴⁹ In this case there is no misidentification, because there is a squeezing which the subject does, but not the one which she sees the anatomically congruent hand doing. Each participant sees the experimenter’s hand where hers is supposed to be. (Probably the participant can misidentify who acted in the ‘I’-thought ‘I performed this action which I am paying attention to’. But this would be the claim that the visually identified event of squeezing is my squeezing. Again, the identification of the event is based on vision.) The participant knows of the squeezing event from the inside and in this event who is acting cannot be misidentified.

Still Lane and Liang might think the ‘I’-thought ‘I did it’ is not IEM relative to ‘I’ in the body-swap illusion. But Shoemaker did not have a stance on this. If this is their target then there is another illusion which might serve their purpose better: the pantomime illusion (Wegner 2005). One subject, A, puts her hands where a different subject B’s arms usually are. B does not see his own arms, just A’s arms straight ahead. The experimenter gives an order “Flex your wrist!” A’s wrist flexes where B’s wrist usually is and participant B reports that he felt that he did it. B’s hand was not even there and he was instructed not to flex it! A case like this might be more appropriate for arguing that a claim such as ‘I did it’

⁴⁸ One might think that self-ascription of agency can be based on vision, but nobody would think that such a judgment is immune relative to the action which is individuated by what one sees.

⁴⁹ This is similar to the standard setup in the Rubber Hand illusion.

based on agent awareness is not immune.⁵⁰ But it is not clear that Shoemaker holds this view.

4. Summary

We have seen that the three empirical counterexamples against immunity do not seem to work. At best they can aspire to be effective against one or another explanation of IEM, but not against the phenomenon of IEM.

⁵⁰ For this argument to work one has to argue that vision can be a constitutive part of agent awareness and that all self-attribution of agency has to be based on agent-awareness.

Chapter 6

On the Identification-freedom Explanation of IEM

Abstract

The standard understanding of immunity to error through misidentification is that IEM is identification-freedom, and there is only one kind of IEM. Although this view is simple and attractive, I will argue that Identification-freedom cannot explain IEM in general.

In part I of this chapter, I will argue that there are two different kinds of IEM. In the course of arguing for these claims, I will distinguish in Evans's work two different explanations of IEM: the Identification-freedom explanation and the Special-basis explanation. Of these explanations of IEM, I will show that the first *only* works for explaining the IEM of demonstrative thoughts, while the second *only* works for explaining the IEM of self-ascriptions. This is because the IEM of self-ascriptions and demonstrative thoughts are of different kinds.

In part II, I will argue that there is a form of identification which lacks the logical form $a=b$ but opens the possibility of misidentification. I will distinguish two different kinds of identification. One only opens the possibility of misidentification for self-ascriptions but not for demonstrative thoughts. This supports my previous result that there are two different kinds of IEM.

Part I

The Two Kinds of IEM:

Evans's Identification-freedoms

1. Preliminaries

One of the most worked out theories of immunity to error through misidentification comes from Evans's groundbreaking work *The Varieties of Reference* (1982). Evans's basic idea is that an immune judgment Fa (B) is not grounded on an identification $a=b$; in that case there is no possibility of misidentification. Call this the **Identification-freedom explanation**. This explanation of IEM has yet to be questioned.

As we will see, Evans's explanation of IEM is not as simple and straightforward as it seems at first and as it is assumed to be. There are several difficulties which we have to face in trying to understand his explanation of immunity to error through misidentification.

The first difficulty arises from trying to understand what Evans's explanation of IEM is. While Evans's work is widely assumed to provide us with *the* explanation of immunity, I will argue that it actually presents us with two different putative explanations of IEM. Evans works out the Identification-freedom explanation for explaining the immunity of demonstrative-thoughts, but he generalizes the explanation to cover the immunity of self-ascriptions. The second difficulty concerns this generalization.

The second difficulty threatens most existing theories of immunity including Evans's. It is supposed that certain self-ascriptions, demonstrative-thoughts, and 'here'-thoughts are immune. So different kinds of thoughts can be immune. Although this claim that more than one kind of thought can be immune is widely accepted, the focus of explanations of IEM is always on the IEM of a particular kind of thought.⁵¹ For example, Shoemaker (1968), Peacocke (2008), Recanati (1993, 2007b, 2012, 2013), Higginbotham (2012), and O'Brien (2007, 2012) in working out their notion of IEM concentrate solely on self-ascriptions,

⁵¹ Wittgenstein only discusses the IEM of self-ascriptions.

but they also accept that at least one other kind of thought is immune.⁵² Evans (1982) primarily focuses on demonstrative thoughts and ‘here’/‘now’-thoughts,⁵³ and simply assumes that his explanation also works for immune ‘I’-thoughts. Campbell (2002) is interested in the IEM of demonstrative thought and self-ascription, but his treatment of the immunity of demonstrative thoughts cannot be easily extended to explain the immunity of self-ascriptions.

Something is missing. What is lacking is support for generalizing the explanation of IEM of one kind of thought to all kinds of cases (‘I’/‘that’/‘here’-thoughts). These philosophers might suppose that working out the explanation of immunity for one kind of thought is the way to arrive at a generalized theory of IEM. This would be a good strategy if there were only one kind of IEM. But this does not seem to be the case – as I will argue.

I turn to discuss Evans’s putatively general explanation of immunity: the Identification-freedom explanation. A general explanation should cover all kinds of IEM, i.e. both first-person and demonstrative IEM.

2. Identification-freedom explanation

2.1. The first version: Factual Identification-freedom (Factual IF)

Evans writes:

When knowledge of the truth of a singular proposition ‘a is F’ can be seen as the result of knowledge of the truth of a pair of propositions, ‘b is F’ (for some distinct Idea, b) and ‘a=b’, I shall say that the knowledge is identification-dependent: it depends (in part) on the second basis proposition [a=b], which I shall call the *identification component*. We might

⁵² Shoemaker has IEM for demonstrative- and ‘I’-thoughts (Shoemaker 1968: 558). Peacocke speaks about IEM of demonstrative-, ‘here’/‘now’-, and ‘I’-thoughts. Recanati considers the IEM of demonstrative- and ‘I’-thoughts.

⁵³ Evans was the first theorist to introduce IEM for ‘here’-thoughts. I will not discuss this issue. First of all, I am not sure that ‘here’-thoughts can be immune. If they were IEM, then they might not exhibit the same kind of immunity as what I discuss here. Second, a use of ‘here’ refers to a location and not to an object in the sense that ‘object’ is used to mean proper object. Another difficulty is that the location might be our solar system, the Earth, a country, a city, or a branch of the tree on which I am sitting. It is like Proteus. So when I call it a location I just try to give a name to something which is defined by the context of utterance and other indexical factors. Thus ‘here’ is very different from the first-person pronoun and from demonstratives. A possible treatment of ‘here’ takes ‘here’ to be a contrast with ‘there’, where both terms are determined in the context. In this contrastive understanding misidentification might be possible.

say that knowledge of the truth of a singular proposition is *identification-free* if it is not identification-dependent. (Evans 1982: 180; my emphasis).

On this picture, a thought Fa is identification-dependent when the knowledge of the proposition depends on the truth (correctness) of an identification-component ($a=b$). A thought is immune iff it is, in fact, **not** identification-dependent. An 'a'-thought (Fa) will be immune relative to a if the knowledge of Fa does not depend on the truth of an identification component (with the logical form: $a=b$). This provides us with the first approximation of the Identification-freedom explanation of immunity.⁵⁴

According to this weak Identification-freedom explanation:

(IF ^{factual}) An 'a'-thought Fa is immune relative to a iff Fa is such that the **attribution of F-ness in fact, is not** dependent on an identification component ($a=b$).

I dub this version of the Identification-freedom explanation **Factual Identification-freedom**.⁵⁵ According to this explanation, if a thought Fa is not grounded on identification then *what (who) is F* cannot be misidentified. To make it simple, Fa is grounded on an identification-component ($a=b$) iff the correctness of the subject's attribution of F -ness to a is dependent on the truth of the identification. As Evans puts this point:

If he knows, or takes himself to know, upon the basis of the operation of the channel, that the property of being red is instantiated, then he may judge 'This is red', and such a judgment is identification-free. (...) Since they do not [*note: here we would have expected 'cannot'*] rest upon an identification, they are *immune to error through misidentification*. (Evans 1982: 182; with my emphasis and parenthetical remark)

⁵⁴ One might hold that Identification-freedom is the basic characterisation of IEM and it is not an explanation of IEM. This would not affect my argument. Identification-freedom seems to explain what opens the possibility of misidentification (being grounded on an identification), thus it better fits the job description of being an explanation than of being a characterisation.

⁵⁵ The label 'factual immunity' comes from Shoemaker. But Shoemaker uses it only for quasi-memory/memory judgments.

For Factual Identification-freedom, what matters is that **the thought does not rest upon an identification**, and **this alone** makes the thought (factually) immune for Evans.⁵⁶

When my thought ‘I have my legs crossed’ is based solely on proprioception, then it is not dependent on an identification component. I cannot misidentify whose legs are crossed, thus my thought is immune. When my self-ascription of ‘having my legs crossed’ is based on vision, looking at my legs (or in the mirror), then my self-ascription is dependent on an identification component. I take my leg to be the one which I see. I take it that I am looking at myself. In this second situation my thought is based on an identification-component.⁵⁷

However, being or not being grounded on an identification-component could be contingent. The key question for factual immunity is whether a judgment **in fact** is grounded on an identification-component which may vary from situation to situation.

3. Special-basis Explanation of IEM

Surprisingly, Evans has a different explanation for the IEM of self-ascriptions. Evans seems to acknowledge that dedicated ways of knowing (e.g. kinaesthesia and proprioception) play a constitutive role in the explanation of the IEM of self-ascriptions (1982: 216, 221). Evans writes: “[E]ach of these ways of gaining knowledge of ourselves [e.g. proprioception, kinaesthesia] give rise to

⁵⁶ To take an example, Morgan (2012) holds that this explanation (which I call Factual IF) is the standard explanation of IEM and by far the best and most powerful one. Morgan, like many others, does not realise that in Evans there is another explanation of IEM (the Special basis explanation).

⁵⁷ Peacocke (2012) seems to think that if I see my leg attached to my torso, then, based on vision, my self-ascription cannot be misidentified. It is not likely that I will misidentify whose legs are crossed, but it is possible. It can seem to me that it is my leg when it is someone else’s leg. It can be an art (or philosophy) installation involving a mirror where someone else’s leg looks to be where mine is. Such cases of misleading appearances are very unlikely in ecologically valid situations. But it does not matter how likely it is that I actually misidentify whose legs I see. What matters is whether misidentification is excluded. In contrast, when I feel pain from the inside and, based on this, I think I am in pain misidentification is not possible. Experiments exploiting the Rubber Hand Illusion paradigm demonstrate that I can feel that a limb is mine when it is not attached to my torso (Knoblich et al. 2005, Longo and Haggard 2012).

judgments which exhibit [...] ‘immunity to error through misidentification’”(Evans 1982: 216).

Evans seems to hold that there are certain special bases, like proprioception and kinaesthesia. Self-ascriptions which are based only on such special bases are immune. This thought suggests another explanation of IEM, which only works for explaining the IEM of self-ascriptions made on a certain basis; this is the Special-basis explanation (already discussed in chapter 4).

Let me call a way of knowing of the subject which is dedicated only for knowing of herself a **dedicated way of knowing**. (I understand the ‘basis’ of a judgement in terms of the relevant way of knowing, Evans uses the notion of ways of knowing as well. I use W for the relevant way of knowing through which the belief/knowledge is acquired; in my terminology $B=W$.) When a self-ascription is based only on a dedicated way of knowing then this is necessary and sufficient for the self-ascription to be immune. Proprioception, nociception, sense of balance, and several other ways of knowing, including introspection, are dedicated ways of knowing. This explains why knowing that I have my legs crossed (W =proprioception), or I am in pain (W =feeling) is immune.⁵⁸

I called this the **Special-basis explanation**. Evans seems to assume that this is only a version of the Identification-freedom explanation. But I argue that this provides a different explanation of IEM from the Factual Identification-freedom explanation. How are they different? First, according to the Special-basis explanation, the basis of the self-ascription is a constitutive part of the explanation of IEM. Second, the Special-basis explanation implies basis-relativity. Third, since being based on a dedicated basis **excludes the possibility that the judgment could involve identification**, this explanation is modally stronger.

The Special-basis explanation can already be found in Shoemaker (chapter 4). However, Shoemaker does not discuss whether a judgment can be immune when it is based on certain forms of perception, like proprioception and kinaesthesia. But Evans tries to argue for this (Evans 1982: 220-221). The

⁵⁸ The basis is the way of knowing $B=W$ according to my interpretation. Still using W seems to be more precise. For me basis-relativity means relativizing the judgment to the way of knowing through which the knowledge is acquired.

Special-basis explanation allows the self-ascription of not only psychological but physical predicates (e.g. having one's legs crossed) to be immune.

What enables the subject to systematically self-ascribe properties known through dedicated ways of knowing? If the subject (tacitly) knows that nothing but herself can be known through a dedicated way of knowing, then it is plausible to think that self-ascriptions based only on such ways of knowing are immune.⁵⁹

However, it seems hopeless to extend the Special-basis explanation to explain the IEM of demonstrative thoughts. For demonstrative-thoughts the way of knowing may be vision or touch or other perceptual faculties, but through all these ways we can know more than one object. This suggests that there might be a particular kind of first-person IEM which is different from demonstrative IEM.

However, Factual Identification-freedom and the Special-basis explanation might only be two different formulations of the same idea. To argue that the explanations are different and that one of them is stronger, I have to show that we can have one of them without the other. Could Factual Identification-freedom accommodate the Special-basis explanation?

3.1. Could the Special-basis explanation be a version of Factual Identification-freedom explanation?

If the Special-basis explanation could be translated into Factual Identification-freedom explanation then they are basically the same explanation. Thus if I want to show that they are different I have to rule out that this can be done.

The difficulty arises from the following fact. For the subject, knowing an object and a property through dedicated ways of knowing makes **misidentification impossible**. For Factual IEM what excludes misidentification is that the subject's knowledge of the judgement does not rely on an identification component – nothing else is required. In the Factual Identification-freedom explanation of IEM the way of knowing is not even

⁵⁹ How can we find evidence for such tacit knowledge? This might be inferred from considering the subject's propensity and stability in following a norm. The norm is that whenever something is known from the inside then it is self-ascribed by the subject. For further discussion of tacit knowledge see Peacocke (1983).

marked. This would be required to accommodate the Special-basis explanation of IEM. Thus I have found an unbridgeable difference between the two explanations of IEM.

It could be suggested that we should put a variable for the relevant way of knowing into the formulation of Factual IF. But simply adding a variable would not make the variable a constitutive part of the explanation. And this is what we need when we ask for the basis.

To show that this unbridgeable gap is real I have to try to show why Factual Identification-freedom cannot accommodate the Special-basis explanation. When I judge based on proprioception then my judgment **in no condition can be grounded on identification**; and we might think that when it is based on a mirror then it **has** to be based on an identification component. Consequently, the real question is how is it possible that an immune self-ascription cannot be dependent (as opposed to: is not dependent) on identification?

My suggestion is that focusing on the distinctive way of knowing helps. The relevant way of knowing is such that on this way of knowing the immune self-ascription **cannot** be dependent on identification. This might explain why for Evans the way of knowing matters for deciding whether a self-ascription is immune.⁶⁰ Another point in favour of this suggestion is that in discussing immune 'I'-thoughts, Evans criticises Wittgenstein for neglecting the basis of the judgment:

Wittgenstein's discussion does not take sufficient account of the fact that the property of immunity to error through misidentification is not one which applies to propositions *simpliciter*, but one which applies only to judgments made upon this or that basis. (Evans 1982: 219).

Why would the basis of a judgment be important for IEM? Evans holds that the subject's ways of knowing (like proprioception) of herself have special properties which exclude that a thought based on them **can** be based on an identification-component (Evans 1982: 180, 220-221). But this is more than

⁶⁰ Evans uses 'way of knowing', 'way of gaining knowledge' and 'channels of information' interchangeably.

what Factual Identification-freedom can accommodate. Thus we need to introduce a new version of Identification-freedom: Strong Identification-freedom. To show the difference between Factual Identification-freedom and the Special-basis explanation, I will show that only a new version of the Identification-freedom explanation can accommodate the Special-basis explanation. This will allow us to clearly see what the relevant differences between the two explanations are.

4. Strong Identification-freedom

For a judgment, Fa (W) it is possible that in one situation on the same basis, W, it is grounded on identification, while in another situation on the same basis, W, it is not grounded on identification. Accordingly, we have to ensure that Fa (W) cannot be grounded on an identification-component if it is immune. To ensure that Fa cannot be grounded on identification it is not sufficient that it is, *in fact*, not grounded on identification – which is only contingent. Rather, something has to exclude the possibility that Fa can be grounded on identification. We need to specify that Fa (W) **cannot** be dependent on an identification-component on this very basis, W. This is the reason why the judgment has to be relativized to its basis for being immune; the basis guarantees that the judgment is immune on this particular basis.

Strong Identification-freedom **excludes** that an immune judgment made on a particular basis (like proprioception) **can be grounded on identification**, in contrast with Factual Identification-freedom. The latter only excludes that the judgment, Fa (W), is in fact grounded on identification. Let me introduce some notation. I will refer to a judgment Fa which is based on a way of knowing W like this: Fa (W). What is absent from factual immunity is the idea that the basis, W, on which the immune judgment is made **cannot be** dependent on an identification-component ($a=b$). Let us recruit this idea.

Let us dub this form of strong Identification-freedom where the basis ensures that the thought cannot be misidentified **Strong IF**:

(IF^{strong}) An 'a'-thought based on W is immune iff the **ascription** of *F* based on W **cannot be** dependent on an identification component ($a=b$).⁶¹

Recall: 'I have my legs crossed' is immune in one but not immune in another situation. The difference between the two situations was the basis of the thought. In the first situation the thought was based on proprioception and in the second case it was based on vision. A suggestion is that, for Strong IEM, the basis makes it the case that **on the relevant basis it is not possible to be grounded on an identification ($a=b$)**. When *Fi* is based only on proprioception, this indeed excludes the possibility that the judgment could involve an identification ($i=b$). Only one object, the subject, can be known through a dedicated way of knowing according to the Special-basis explanation. This makes the basis, the way of knowing, constitutive in the explanation of the IEM of self-ascriptions. And this provides good reason to include a variable for the basis, the way of knowing, of the thought in the explanation of the IEM of self-ascriptions. This suggests that there is a special form of first-person IEM (strong IEM) which is different from factual IEM. Thus, we see that Strong Identification-freedom provides a way of accommodating the Special-basis explanation.

Let me make it clear how I use the notion of 'basis'. I speak about relativizing the judgement to its basis. By using 'basis', I only mean **the way of knowing**, like proprioception, introspection or vision, which is employed in forming the relevant judgement. But this is meant to be different from the ground of justification for the judgement. Shoemaker's use is similar to mine when he discusses special way of knowing, and Evans seems to endorse a similar idea that *only* the way of knowing makes a difference for first person IEM.⁶² Following them, I distinguish between the notion of 'way of knowing' and

⁶¹ It is important to note that 'cannot be' denotes the denial of epistemological possibility. Epistemological possibility requires that something is possible, for all the subject knows (Edgington 2004). My question is not whether the possibility under consideration is physically or logically possible but only whether it is epistemically possible. Given what the subject knows, it cannot be the case that the self-ascription is dependent on an identification component.

⁶² "[I]t is necessarily the case that if a person knows that a P*-predicate applies to him he knows that it applies to him in the "special way" appropriate to that predicate (which does not preclude that he should also be in a position to know that it applies to him in other ways, i.e., ways in which others might know that it applies to him). A more explicit formulation is this: ϕ is

the notion of ‘ground of justification’. ‘Ground of justification’ means whatever puts the subject in a position to make the judgement in question. The ground of justification involves more than the way in which knowledge is gained. It involves details of the situation which enable the subject to make the judgement as well as the way of knowing employed.

Once I make this distinction, a difference between first-person IEM and demonstrative IEM becomes clear. What matters for first person IEM (Strong Identification-freedom) is only the way of knowing through which the knowledge (or belief) is gained. For demonstrative IEM (factual IEM) the ground of justification matters. It has to be settled whether there is in fact identification involved, since this cannot be guaranteed just from examining the basis (way of knowing) of a judgement. For factual IEM what matters is only whether or not the knowledge of the judgement involves an identification.⁶³

One might worry that strong immunity is what Shoemaker calls absolute immunity. Absolutely immune self-ascription cannot be misidentified relative to ‘I’ in any circumstance and on all bases. In contrast, according to strong immunity the judgment is not immune *per se*, but is immune on a particular basis W, but not on all bases. Strong Identification-freedom presupposes basis-relativity. Accordingly, Strong Identification-freedom is not absolute IEM.

To sum up, *according to Strong IF, an ‘a’-thought (W) is immune iff the ascription of F-ness **cannot be** dependent on an identification component with the logical form (a=b).* The basis, W, is such that a judgment made on that basis, W, cannot be dependent on an identification component. In contrast, factual immunity merely requires that the thought, *Fa*, is *in fact* not dependent on an identification component (*a=b*).

I want to suggest that Strong and Factual Identification-freedom explain different kinds of IEM. First, the basis of the self-ascription is a constitutive part of the explanation of IEM in Strong but not in Factual Identification-freedom.

a P*-predicate if and only if there is a way *w* of knowing *ϕ* to be instantiated such that, necessarily, S knows *ϕ* to be instantiated in way *w* if and only if S knows that he himself is *ϕ*.” (Shoemaker 1968: 565). “[W]e cannot think of the kinaesthetic and proprioceptive system as gaining *knowledge* of truths about the condition of a body which leaves the question of the identity of the body open” (Evans 1982: 221).

⁶³ Note that the notion of *ground of justification* can only be assigned to a token judgement while if we consider a judgement *relativized to the relevant way of knowing* (e.g.: proprioception) we are picking out a type of judgement subsuming a range of tokens.

Second, Strong Identification-freedom implies basis-relativity, but Factual Identification-freedom does not (where ‘basis’ means way of knowing). Third, Strong Identification-freedom is modally stronger than Factual Identification-freedom. A judgment can be factually immune when it is not strongly immune. Strong Identification-freedom seems to work perfectly well for explaining the IEM of self-ascriptions. This shows that there is a particular kind of first-person IEM. But could Strong Identification-freedom work for explaining demonstrative IEM?

5. The Strong IF explanation is Hopeless for Explaining the Immunity of Demonstrative Thoughts

Consider a case when I look at a blue tit. Suppose I do not think that I have seen that blue tit before and I do not think of that bird in any other way besides looking at it. It is very busy eating. I judge: ‘That one is greedy.’ Is it possible that I misidentify *what is greedy*? If I do not think or assume that I have seen this tit before, then, on this basis, which I have, I cannot misidentify what is greedy. In this case ‘That one is greedy’ is immune relative to ‘that one’ in it.⁶⁴ Does this basis, vision, exclude the possibility of misidentification?

It is possible to misidentify which particular one is greedy based only on vision. Suppose that I see a greedy tit at t_0 , turn my head away, and, when I turn my head back, I see a tit where the greedy tit was at t_1 . I can misidentify which particular tit is greedy. I may attribute being greedy to the tit which I see at t_1 because I assume or think it is the same one which I saw at t_0 . I might say of the tit I see at t_1 (which is not eating at the moment) that this one is greedy, even though it is a different tit from the one which I saw and was greedy at t_0 .

⁶⁴ If there would be a sudden substitution of greedy tits, then according to Evans I could not refer to it. In order to refer by using a demonstrative, it is necessary that there is one object to which I attend to and to which I intend to refer to. If I try to refer to two objects with a singular demonstrative referring expression, then I fail to refer in his view. Evans allows that I might have another kind of thought, like a general thought, in this situation, but I cannot have a demonstrative thought of a particular object. This is because there is no particular object to which I intend to refer to. Thus when I succeed in referring, it is guaranteed that whatever I know should be about that single object to which I intend to refer to. On other views demonstratives might fail to refer and successful reference has certain conditions. On such views IEM is only for such demonstrative thoughts which involve **referring** demonstrative terms.

The Strong Identification-freedom explanation requires that the basis of the thought has to be such that if *Fa* is made on such basis, then *Fa* cannot be grounded on identification. But I would be in trouble if I have to find a demonstrative thought which is based on vision, and where it cannot be grounded on identification just because it is based on vision.

Recall: if I judge that I have my legs crossed based only on proprioception then on this basis it is not possible that I misidentify who has her legs crossed. To use Evans's words: "[W]e cannot think of the kinaesthetic and proprioceptive system as gaining *knowledge* of truths about the condition of a body which leaves the question of the identity of the body open" (Evans 1982: 221). But vision is not like that; based on vision, nothing excludes the possibility that misidentification **can** happen.

Surprisingly, Factual Identification-freedom can explain the IEM of demonstrative thoughts perfectly. When a demonstrative thought based on vision is grounded on an identification component then it is open to misidentification. For 'that is *F*', when I have an earlier encounter through vision with the assumed-to-be-identical object, and the attribution is based on this earlier encounter, then I can misidentify which object is *F*, since the object may look like another object. Thus being based on vision alone cannot exclude the possibility of misidentification, unlike being based on proprioception or introspection. This suggests that there is a different kind of IEM: demonstrative IEM.

Strong Identification-freedom requires that the thought based on a way of knowing *W* cannot be grounded on identification. Consequently, Strong Identification-freedom cannot explain the IEM of demonstrative thoughts, although it seems to work nicely for explaining the IEM of self-ascriptions. The ways of knowing on which perceptual demonstratives can be based are not like proprioception. Vision or touch, on which demonstrative-thought can be based, do not exclude the possibility of identification. Thus, Strong Identification-freedom is hopeless to explain the immunity of demonstrative thoughts.

It seems that we have two different kinds of IEM explained in two different ways for two different kinds of judgements (self-ascription and demonstrative judgement). There are different conditions for what excludes misidentification

for self-ascription and for demonstrative thought. These are the reasons why we have to distinguish first-person IEM and demonstrative IEM. These are two different kinds of IEM.

6. General Explanation of Immunity

It is widely held that immunity is a unified feature because what explains immunity is supposed to explain all instances of it. This is unproblematic unless we face more than one kind of IEM. When we find an explanation of why all kinds of thought (first-person, demonstrative), which are immune, are immune, then this provides a ***general explanation of immunity***. But it will be much harder to provide a general explanation of IEM if we find that there is more than one kind of IEM – as I showed.

Factual Identification-freedom only explains the IEM of perceptual demonstrative thoughts, while Strong Identification-freedom seems to work only for immune self-ascription. Thus neither explanation provides a general explanation of immunity of all kinds of thoughts. The conditions of first-person and demonstrative IEM are different, so if there is a general explanation of IEM, it needs to leave room for these differences.

One might suggest that Factual Identification-freedom is compatible with Strong Identification-freedom. Thus Factual Identification-freedom should be the general explanation of IEM. The problem with this suggestion is this. The condition which excludes misidentification is different for self-ascriptions and for demonstrative thoughts. The explanation of IEM should explain what excludes the possibility of misidentification. However, Factual Identification-freedom only explains this for immune demonstrative thoughts but not for immune self-ascriptions. Conversely, Strong Identification-freedom cannot be the general explanation of IEM because it cannot explain the immunity of demonstrative thoughts.

It would appear that Evans is engaging in the project of finding a general explanation of IEM, but it seems that neither kind of Identification-freedom works for that purpose.

6.1. The Absent General Explanation of IEM

We have not seen a general explanation of IEM yet. I will turn to this in the next chapter. If an explanation explains only one kind of IEM but not all kinds of IEM then this would be **a special explanation** of this kind of IEM. We have only found in Evans two special explanations of IEM for the two different kinds of IEM. Factual Identification-freedom is the special explanation for demonstrative IEM and Strong Identification-freedom (or the Special-basis explanation) is the special explanation for first-person thought.

7. Summary for Part I

I distinguished two kinds of IEM: first-person IEM and demonstrative IEM. The difference is cashed out in terms of the difference between Factual and Strong Identification-freedom. According to the Strong IF explanation, Fa (W) is strongly immune iff the attribution of F -ness based on W **cannot** be grounded on identification ($a=b$). The modally weaker version of Identification-freedom is called Factual IF. According to the Factual Identification-freedom explanation, Fb is factually immune iff the attribution of F , *in fact*, **is not** grounded on identification.

Part II

Varieties of identification:

Two kinds of identification

1. Preliminaries

The consensus about IEM is that it just is identification-freedom.⁶⁵ I have shown there are too many explanations in terms of Identification-freedom, but this is not the only problem. I will try to show that there is an ambiguity in the notion of ‘identification’.

⁶⁵ Recanati 2007, 2012, O’Brien 2007, Peacocke 2008: 279-280, 2008/2009, Peacocke 1983. Note that for Peacocke, immune introspection-based self-ascription is based on an a priori true identification ‘I = the thinker of this thought’. But otherwise identification-freedom holds.

Identification-freedom is designed to explain IEM in general.⁶⁶ Identification opens the possibility of misidentification.⁶⁷ According to Factual Identification-freedom, an 'a'-judgment *J* based on *W* is immune relative to 'a' iff *J* based on *W* is in fact not grounded on an identification ($a=b$). To make it simple, a judgment (Fa) is grounded on an identification ($a=b$) or presupposes this identification if the subject's knowledge of Fb is dependent on the truth of that identification.

2. Constraints on the satisfactory explanation of IEM

In the immunity literature there is an implicit supposition that there should be a general explanation of IEM.⁶⁸ Let me examine how one usually arrives at a general explanation of IEM.

Consider Peacocke as an example. Peacocke (2008) introduces IEM with an example 'This keyboard is black', based on looking at a keyboard. Peacocke characterises IEM and generalises what he found about the IEM of demonstratives to 'now'-thoughts (e.g. 'It is raining now') and to the IEM of self-ascriptions (Peacocke 2008: 92-93).⁶⁹ The strategy seems to be to explicate what IEM is for one kind of thought and generalise it to understand the IEM of other kinds of thoughts. As we have seen, this strategy is widely accepted to be correct. Evans used the same strategy, starting with the IEM of demonstrative thoughts and generalising it to other kinds of thoughts (Evans 1982: 180). So it seems that there is a presumption that there is a general explanation of IEM

⁶⁶ All the difficulties for Factual Identification-freedom apply to Strong Identification-freedom, since the latter entails the former. So I will only discuss Factual Identification-freedom here.

⁶⁷ There is more than one kind of misidentification which is discussed in the literature. A careful discussion of these issues can be found in Coliva 2006. I restrict my discussion to present tense immune self-ascriptions and perceptual demonstrative judgments; Coliva's discussion is about self-ascriptions based on memory. I do not discuss past tense self-ascriptions. I take Evans to be discussing circumstantial IEM, although nothing hangs on the acceptance of this supposition.

⁶⁸ This is presupposed in Shoemaker (1968), Evans (1982), Campbell (1999), Coliva (2006), O'Brien (2007), Peacocke (2008), and Recanati (2007, 2013).

⁶⁹ Peacocke assumes that the explanation of IEM for Fx has to rely on the fundamental reference rule of 'x' (Peacocke 2008: 92). Interestingly, Peacocke explicitly mentions only that the IEM of self-ascriptions has to be explained by drawing upon the fundamental reference rule and assumes this for other kinds of immune thoughts, like immune demonstratives. But this seems to be a different explanation from identification-freedom which he endorses. If the fundamental reference rule plays a crucial role in the explanation of IEM (with which we agree), then it is not clear that the explanation of 'I'-judgment and 'that'-judgment will be the same because their reference is fixed differently. Thus what excludes the possibility of IEM might be different for different kinds of judgments.

that equally covers all kinds of IEM. Accordingly, one can arrive at this general understanding just by picking one's favourite case of IEM and generalising the result to other cases. Both Evans and Peacocke employ this strategy, like many others. This would be justified if there would only be one kind of IEM and not two. But we have seen that there are two kinds of IEM (first-person and demonstrative).

A general explanation of IEM has to explain the IEM of all kinds of immune judgements. Call this the **Explanation-generality Constraint**. Identification-freedom is designed to explain IEM in general, thus it should satisfy this Constraint.

3. The question

When one starts to think about Identification-freedom one of the first questions to come into one's mind is: **What is the notion of 'identification' involved in identification-freedom?** This answer is necessary for understanding identification-freedom properly.

Initially, we might think 'identification' should be considered in terms of an identity statement with the logical form: $a=b$. Evans clearly claims that the identification component has the logical form $a=b$ – exactly what we would expect. I will first provide examples of two kinds of identification and then I turn to argue that we should treat them as different kinds of identification. I turn to discuss the first notion from Evans.

4. 'Identification' means Frege-identification

When we consider an identification of the form $a=b$, it is best understood following Frege's analysis. I have to give a quick and simplified recapitulation of Frege's famous puzzle to set up our understanding of what we call Frege-identification. Suppose 'a' and 'b' are singular terms referring to an object. Frege recognised a difference between (1) $a=b$ and (2) $a=a$.

4.1. The Frege's puzzle

If two identifications have the same cognitive significance then the subject's cognitive relation to those two relations will be the same. But a subject may

doubt that $a=b$ even if she would never doubt that $a=a$. The difference cannot be accounted for only by relying on what is the referent of the two relata. The two identity claims ($a=a$, $a=b$) mean the same in terms of reference: the referent is the same. However, the discovery that Hesperus is Phosphorus has shaken our common knowledge, but nobody can discover what everyone knows – that Hesperus is Hesperus. Frege observed a difference between $a=a$ and $a=b$ in their cognitive significance which needs to be explained. This is called **Frege's puzzle**.

In "On Sense and Reference", Frege introduces the notion of sense to explain the difference between $a=b$ and $a=a$. The reason why $a=b$ can be a discovery is because 'a' and 'b' have different senses, while 'a' has only **one** sense in $a=a$. Frege understands sense in terms of modes of presentation (hence MPs) which account for the cognitive difference between $a_{MP1}=a_{MP1}$ and $a_{MP1}=b_{MP2}$. 'Hesperus' and 'Phosphorus' have different modes of presentation in the form of different definite descriptions. An MP might be a definite description like 'the last visible heavenly body to disappear at dusk' for Phosphorus. But an MP need not be a definite description. The introduction of different modes of presentation (MPs) for $a=b$ solves Frege's puzzle.

Frege is often taken as claiming that the sense (MP) of a singular term fixes what the referent of that term is.⁷⁰ According to Frege, given that the mode of presentation (the way that the reference is fixed) is the same, then the subject cannot doubt that $a_{MP1}=a_{MP1}$. Thus, self-identification is not a Frege-identification. Frege-identification *requires* that the two singular terms have to have different senses – MPs – ways of fixing the referent. Consider the figure below illustrating the difference.

⁷⁰ "The regular connection between a sign [=singular term], its sense, and its reference is of such a kind that to the sign there corresponds a definite sense and to that in turn a definite reference, while to a given referent (an object) there does not belong only a single sign." (Frege 1948: 564).

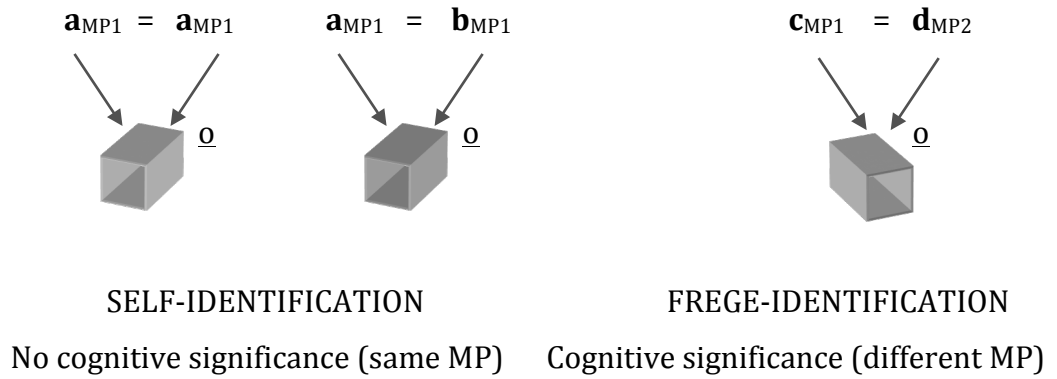


Figure 1. Self-identification and Frege-identification.

Recapitulating, **Frege-identification** (sometimes: **FI**) requires that there should be two referring terms and that the modes of presentation of the two different terms have to be different: ' $c_{MP1}=d_{MP2}$ '.

For Evans, immunity consists in identification-freedom. Accordingly, identification-freedom states that if the subject's knowledge of a judgment made on a certain basis B rests on a Frege-identification then this judgment based on B is not immune.⁷¹

4.2. Evans uses Frege-identification in the identification-freedom explanation of IEM

Evans seems to be claiming that 'identification' in identification-freedom is simply Frege-identification. An 'a'-judgment (B) is identification-free relative to 'a' iff, based on B, it is *in fact* not Frege-identification dependent relative to 'a'. In Evans's account, a judgment '*a* is *F*' (B) is identification-dependent when it can be seen as a result of a pair of judgments

- (1) '*b* is *F*'
- (2) ' $a_{MP1}=b_{MP2}$ ' (this component is called the 'identification component').⁷²

⁷¹ Evans distinguishes between IEM in the narrow and in the broad sense. Here I only discuss IEM in the narrow sense, because IEM in the broad sense works only for descriptive names which are introduced into the language by a definite description (e.g. the 'Julius' case). I discuss IEM in the broad sense in the next chapter.

⁷² "When knowledge of the truth of a singular proposition '*a* is *F*' can be seen as the result of knowledge of the truth of a pair of proposition, '*b* is *F*' (for some distinct Idea, *b*) and '*a=b*', I shall say that the knowledge is identification-dependent: it depends (in part) on the second basis proposition, which I shall call the identification component. We might say that knowledge of the truth of a singular proposition is identification-free if it is not identification-dependent. (...) Since they do not rest upon an identification, they are *immune to error through misidentification*." (Evans 1982: 180-182).

An 'a'-judgment is identification-free iff it is not Frege-identification dependent. An immune judgment $J(B)$ is identification-free, because $J(B)$ it is not grounded explicitly or implicitly on an identification component. When a judgment, $Fb(B)$, is grounded implicitly on an identification component, then the presupposition of the judgment contains an identification component ($a_{MP1}=b_{MP2}$), and the attribution of F -ness is based on $a_{MP1}=b_{MP2}$.⁷³

But is it the case for Evans that the mode of presentation has to be different? Evans requires that the idea of 'a' and the idea 'b' have to be different in $a=b$ (Evans 1982: 180-182). For Evans, an 'Idea' of a singular term means, roughly, the rule which fixes the referent of a singular term to a unique object. As Frege understood it 'the sense/MP of a singular term' is what fixes the referent of that term to an object. Accordingly, for Evans, in $a=b$, the MPs (Ideas) of the two terms have to be different. Consequently, $a=b$ is a Frege-identification ($a^{MP1}=b^{MP2}$) for Evans.

Shoemaker, in his book (1996), uses the notion of identification-freedom so he seems to agree with Evans that IEM just is identification-freedom. I will try to elaborate a point of disagreement which has not been realized by either of the two parties (Evans or Shoemaker). 'Identification' has different meanings for them.⁷⁴

5. 'Identification' means object-identification

There is a second notion of identification: object-identification. **Object-identification (OI)** involves automatic and involuntary recognition of a

⁷³ Without the last clause, we can easily have a counterexample to identification-freedom. Suppose a subject thinks

1. I am Wittgenstein. (B=testimony, I had amnesia and my friends persuaded me that I am Wittgenstein.)
2. Wittgenstein only has headaches before it rains. (I read this in a biography which ends before Wittgenstein passes away.)
3. I have a headache.

So I conclude that I think it is going to rain. Although my judgment is grounded on an identification, the attribution of the property 'thinking that it is going to rain' is not grounded on the identification. So it is not the case that whenever a judgment is grounded on an identification it loses its IEM.

⁷⁴ Although Evans claims that Shoemaker does not realise that immunity just is identification-freedom, Shoemaker disagrees. Shoemaker (1996) even switched from using the name *immunity to error through misidentification*, which he coined for the phenomenon in question, to *identification-freedom*.

perceptually observed object as *a*.⁷⁵ An example of object-identification is when we see a friend on the street. I know Hildegard well and I see Hildegard pass by on a bicycle. Recognition of Hildegard can be very quick, involuntary and automatic. This is underpinned by a subpersonal mechanism enabling me to directly see an object as an individual like Hildegard (such as the posited face-recognition module located at the fusiform face area of the brain; Kanwisher et al. 1997). Similarly, I can see and automatically and involuntarily recognise my mother, or familiar objects like the sun or the moon. I call this kind of automatic and involuntary recognition of objects **object-recognition**.

The kind of recognition which is object-recognition is similar to automatic and involuntary kind-recognition. One way of automatically recognising redness, being a lion or being a pumpkin is just to see them directly as such. As one sees the redness of the object, one sees Hildegard as Hildegard directly.

Object-recognition happens at the level of perception. The perceptual system is different from the belief system. If I believe *p* and not *p* this is a contradiction. In contrast, I may experience that *p* while I know that *p* is not the case. This is not a contradiction. Contradiction would require that both *p* and not *p* are in the belief system. This kind of object-recognition of individuals happens at the level of perception. One might lack this perceptual mechanism – in prosopagnosia – and have difficulty even in recognising one's mother or brother.

I may see a table as a table, just as I see Hildegard as Hildegard. Seeing a table as a table does not require that I presuppose that the thing over there is a table; it does not presuppose a background belief that this is a physical object. What is missing if I do not see a table as a table is not a presupposition but merely automatic recognition of its kind. Similarly, for individuals one is acquainted with, one simply recognises them automatically (in usual circumstances). I recognise Hildegard as Hildegard. I see Hildegard when I see her as falling under the individual concept 'Hildegard'. Object-recognition can be understood as the capacity to automatically and involuntarily see something as something falling under an individual kind. My perception is different when I see Hildegard as Hildegard automatically. Seeing Hildegard automatically as

⁷⁵ This is a different notion of recognition from Wittgenstein's.

Hildegard is a different experience than not seeing her as Hildegard, and also different when I am unsure whether the person I see is Hildegard. Object-recognition is automatic, and as such does not require a demonstrative identification like 'this one is Hildegard' or 'this one is my mother' – neither as a thought nor as a presupposition.

5.1. 'Identification' means object-identification for Shoemaker

For Shoemaker (and probably Wittgenstein) identification is identification of an object as *a*. I want to suggest that Shoemaker may have understood this as involving object-recognition. If I think *a* is *F*, because I perceive the object which I know to be *F* as *a*, then my judgment *Fa* is open to misidentification. As Shoemaker puts this point:

[T]his is perhaps the most important point, the identification of a presented object as oneself would have to go together with the possibility of misidentification, and it is precisely the absence of this possibility that characterizes the use of 'I' that concerns us. (Shoemaker 1968: 561, my emphasis).

If I say "I am bleeding," it can happen that what I say is false even though I am giving expression to the knowledge that a certain person is bleeding; it may be that (...) I am mistaken in thinking the person who is bleeding to be myself. Such statements are subject to error through misidentification relative to the first-person pronouns, where to say that a statement "a is Φ " is subject to error through misidentification relative to the term 'a' means that the following is possible: the speaker knows some particular thing to be Φ , but makes the mistake of asserting "a is Φ " because, and only because, he mistakenly thinks that the thing he knows to be Φ is what 'a' refers to. (Shoemaker 1968: 556-557, my emphasis).

For Shoemaker, identification is when I see an object as myself (1968: 561), which means, for the subject, that he "thinks that the thing he knows to be *F* is what 'a' refers to" (557). Here we do not have an identification but rather a

referring term 'a' and its referent. This only requires automatic and involuntary recognition of *a*. Thus, Shoemaker means object-identification by 'identification'; he, at least, allows object-identification to open the possibility of misidentification.

When I see my limbs then (usually) I automatically and involuntarily recognise myself. This opens the possibility of misidentification, because the one whom I observe (e.g. with a bleeding hand) might not be me (Shoemaker 1968: 556). But this does not necessarily presuppose an identification component: 'that one is me', or something similar.

Frege-identification has the logical form $a=b$, but in object-identification the relation between the referent and the referring expression is not identity. Object-identification requires an object to be perceptually present and a referring expression which the subject uses for that object as its referent. For object-identification, an object has to be present to me through perception, but not by merely thinking about that object.⁷⁶

Object-identification is neither an identification of *a* with *b* or a presupposition of it. Object-identification involves automatically and involuntarily perceiving an object as *a* (seeing it as falling under an individual concept). For that we need not have two referring terms in mind as pointing to a single object or a presupposition of an identity statement. This seems different from Frege-identification. (I will argue below that OI and FI are different kinds of identification.) Thus Shoemaker means at least object-identification by identification and holds that object-identification opens the possibility of misidentification, contrary to Evans.⁷⁷ Accordingly, Frege-identification is not the only kind of identification which opens the possibility of misidentification, as Evans assumed.

To recapitulate, I supposed with Evans that 'identification' simply means Frege-identification. But as Shoemaker observed, a non-immune judgment can be based on object-identification without being based on Frege-identification.

⁷⁶ By 'object', I mean something like an entity which has some kind of spatial boundary, unity (falling under a kind category like person, house or eggplant, etc.), persistence conditions, and which can be perceived.

⁷⁷ This allows that both Frege-identification and object-identification open the possibility of misidentification, since Shoemaker does not deny that Frege-identification opens the possibility of misidentification.

Thus, 'identification' cannot simply mean Frege-identification as Evans suggested.

In the rest of this chapter I will do the following. First, I will present some examples for judgments involving object-identification and Frege-identification. Second, I will argue for the claim that object-identification and Frege-identification are different kind of identifications. Third, I will argue that not all cases of object-identification open the possibility of misidentification.

6. Frege-identification and object-identification

6.1. Examples for the difference between OI and FI

The following judgments or descriptions of experiences presuppose object-identification, or Frege-identification. A diagram below charts which thoughts presuppose Frege-identification and which thoughts presuppose object-identification.

A: I see Bob as Bob. (I see someone as Bob – falling under the individual concept 'Bob' provided by my capacity of individual face-recognition.)

B: Tully wrote the Philippics. (I read this about Cicero and I know that Cicero is Tully.)

C: I see that Big Ben is there. (I come to believe it after careful consideration about what this big clock tower could be.)

D: I see Big Ben. (I am acquainted with it and am able to recognise it instantaneously. I see it as Big Ben – as falling under this individual concept.)

E: We saw that bird (which we see now) in the morning and called it 'Pavarotti'. (This is based on many independent clues: it looks the same and it sings and eats in the same way. This presupposes that the bird I see now is the same bird I saw before.)

F: That necktie (which I touch but do not see) is red. (I remember putting a red one under the table yesterday. This is Shoemaker's example.)

G: That dog is barking. (This is based on hearing and seeing a dog which looks like it is barking.)

H: I see my mother. (I recognise her without inference. I see her as my mother.)

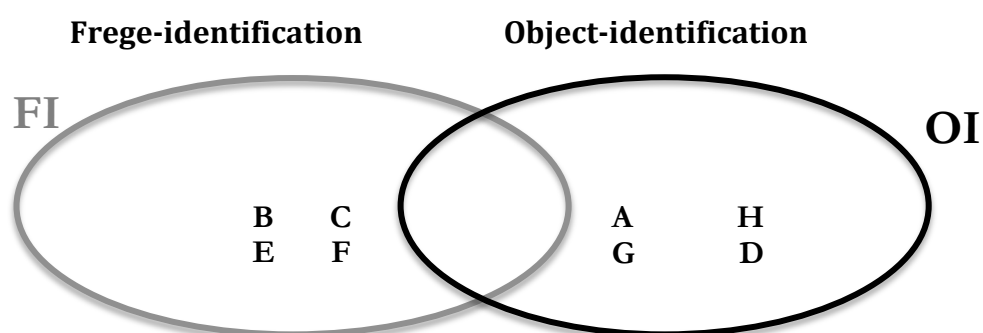


Figure 2. Varieties of identifications: the examples.

The diagram presents cases of two kinds of identification (FI, OI). We found cases where Frege-identification is involved but object-identification is not involved and *vice versa*. Thus, again we find additional support for the claim that being based on object-identification is different from being based on Frege-identification.

One might still worry that there is no real difference between Frege- and object-identification. Either they are identical ($FI=OI$), or OI might be a subset of FI or FI is presupposed by OI. Another harmless option would be that both Frege-identification and object-identification open the possibility of misidentification. In this case ‘identification’ is a name covering both kinds of identification which I distinguished. I will argue for two claims, (i) object-identification is different from Frege-identification ($a_{MP1}=b_{MP2}$), and (ii) it is not the case that both (OI, FI) always open the possibility of misidentification.

6.2. The differences between Frege- and object-identification

To make the difference between the two kinds of identification clear, let us look at a schematic diagram. We simply can read off the differences between the two kinds of identification.

(In the diagram o stands for an object and not for the name of the object.)

The schematic difference is the following:

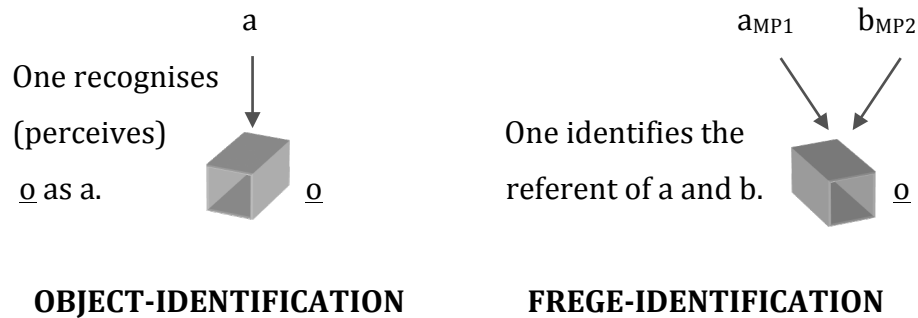


Figure 3. Object-identification and Frege-identification.

Frege-identification is of the form $a=b$. (I will use the symbol \Leftrightarrow for logical equivalence, and will use \Rightarrow for logical implication.) Frege-identification is a symmetric ($a=b \Leftrightarrow b=a$), transitive ($((a=b), (b=c) \Rightarrow (a=c))$) and reflexive ($a=a$) relation.

Object-identification is a relation between a name/singular referring term 'a' and an object when it is based on object-recognition. One takes 'a' to refer to that object. One sees the object automatically and involuntarily as a . When I use 'a' to refer to an object, then I cannot use o (the referent) to refer to a . Why? It is because the object is in the world as an object and it is different from a singular referring term which is in the language. Thus they cannot take each other's place. Consequently object-identification is asymmetric. Object-identification is not reflexive because I cannot take a to be the referent of a . Object-identification is not a version or a subset of Frege-identification because it is not symmetric, not reflexive and not transitive. Frege-identification is an equivalence relation and object-identification cannot be. If object-identification is taken to be a subclass of Frege-identification, then it would have to specify an equivalence relation as well. But it cannot be an equivalence relation because it is neither symmetric nor reflexive. Thus, object-identification can neither be a subclass of Frege-identification nor can it be equivalent to Frege-identification.

An opponent might say, accepting the difference between the two relations, that we do not need object-identification because Frege-identification is all we need for thinking about reference. My response is that this is implausible, because the ability to make a Frege-identification presupposes the capacity of object-identification: the capacity to see an object as falling under a concept or individual categorization. If one cannot track and recognise objects which one

knows at all then one cannot use a singular term to refer to them. When one refers to an object, one uses a referring expression to refer to that object. If one cannot recognise (in experience) an object as the referent of a singular term then one does not understand what it means to refer to an object by using a singular term. The capacity of object-recognition is required for the ability to refer.

To conclude, Frege-identification is different from object-identification. Frege-identification is a symmetric ($a=b \Leftrightarrow b=a$), transitive ($((a=b), (b=c)) \Rightarrow (a=c)$) and reflexive ($a=a$) relation but object-identification is none of these. Thus, object-identification can neither be a subclass of Frege-identification nor can it be equivalent to Frege-identification.

6.3. Object-identification is possible without Frege-identification

We employ object-identification (automatic categorisation under an individual concept) all the time. We simply see a flower as a flower, a top hat as a top hat, and a banana as a banana. I need not presuppose 'that one is a flower/a top hat' and so on. When one looks at her mother, one does not think, presuppose or entertain 'that one is my mother'. Usually we do not think, presuppose or entertain judgments like that. There would be too many judgments to entertain. Similarly, I can just simply see myself as myself. You may ask why? When I look at my hand, I need not think, presuppose or entertain 'this is my hand' ($a=b$) in order to judge something about myself (e.g. 'I have a bleeding hand').

In cases of hesitation, it is plausible to think that this requires presupposing and doubting a Frege-identification. This is because if automatic object-recognition is not in play, then one needs to demonstrate the object and then compare different putative identifications. For example, one might wonder whether the person whom one sees is Hildegard or Alexandra. Is that person Alexandra? This requires Frege-identification. But in many everyday situations automatic object-recognition does not require this.

Evans and followers might think that I always have to entertain or presuppose an identification ($a=b$) in the ground of my thought Fa , if misidentification is possible. Their claim is that this is what underpins the truth of the conditional that if it turns out that a is not F , then I still can retreat to the

claim that something else is *F*. But if the only reason is that a judgment (*W*) has to be based on an identification component whenever misidentification is possible, then this is not a sufficient reason as it begs the question.

An opponent of mine would say that there are visual demonstratives/indices and that they are required when one sees *x* as *x*. I can reply as follows. First of all, if there is some demonstrative indexing like a visual demonstrative this is not sufficient to ensure that we have a Frege-identification. Visual indices in this sense (Pylyshyn's (2007) 'fingers of attention' or Peacocke's (2014) notion of visual index) are mechanisms which enable the subject to track an object in the visual field. They are categorically different from demonstratives or referring expressions. Visual indices are a basic feature of the functioning of the visual system and independent of linguistic capacities. Visual indices enable us to track objects but they are not referring expressions in a language. Visual indices or demonstratives are language independent while Frege-identification requires a particular language and not only the capacity to perceive. Furthermore, the mark of referring expressions is that the referent and its sense can be differentiated. But the mode of presentation and referent cannot be differentiated for visual indices.

To sum up, there are no grounds for thinking that there must be a presupposition of the form $a=b$ for object-identification.

6.4. How is object-identification possible without Frege-identification?

If I see that a lion is approaching or the apple is yellow, I, thereby, know that a lion is approaching or the apple is yellow (Cassam 2007). Similarly, when I see Bob as Bob, I, thereby, know of Bob (Dretske 1969). The experience is simply the experience of Bob. A key difference between object-identification and Frege-identification is that object-identification is at the level of perceptual experience (which can be conceptually laden) and Frege-identification is at the level of beliefs.

A key difference between perceptual experience and belief is the following. Beliefs (as opposed to experiences) are always open to reason. In case one encounters evidence against a belief that one has, one is rationally obliged to alter the belief in response to the evidence against it. In contrast,

experiences like visual experiences are belief independent. For example, when one experiences the Müller-Lyer illusion, one will experience two lines of unequal length even though one knows that they are of the same length. Countervailing evidence would not affect one's ability to see an object as falling under a kind. Even if I know that something is green I can still see it as red if the lighting is such as to make it look red. Seeing a *trompe l'oeil* does not make me think that it is three dimensional, and knowing that it is two dimensional does not make me see it as two dimensional. Perceptual experience can present things in a certain way even though we know that things are not that way. This cannot be with belief.

When I presuppose something then when the presupposition turns out to be wrong I ought to revise my belief. But if I am experiencing an object as a house then it might turn out that even when I know that it is not a house I still experience it as a house. It seems belief independent to some degree. Similarly, seeing someone as Bob or only seeing someone without seeing him as Bob is a different experience. I might know that he is not Bob but I still can see him as Bob if the resemblance is overwhelming. It is not the case that it is a mere presupposition which I ought to revise if I am rational or that I am irrational in seeing him still as Bob. Consequently, recognition might only be the automatic perception of a particular object as *a* (falling under an individual concept) without relying on an identity presupposition.

6.5. Misidentification without Frege-identification

Recall, I argued that when one looks at her mother, one does not think, presuppose or entertain 'that one is my mother'. How is that possible?

Recognition as a capacity is acquired before the capacity to use or understand language. We recognize faces instantly early on. A child can recognise its mother without having demonstrative concepts and language. Face recognition is a pre-linguistic capacity that enables us to use names for persons. This capacity does not seem to require that I can entertain demonstratives. The capacity to recognise people allows for the possibility of misrecognition even before language use. We can use names partly because we can recognise the bearer when we see him or her again. Recognition is required

for naming an object or using a demonstrative for it. Object-identification exploits this primitive capacity.

There are people who cannot recognise faces, prosopagnosics. But object-identification does not require face-recognition; this is only one mechanism for it. Another capacity like voice recognition could enable the subject to recognise someone as n (where ' n ' is a proper name). But someone who lacks both the capacity of face and voice (or other forms of) recognition might require identification for being aware that someone observed falls under an individual concept. For example, Oliver Sacks (2010) cannot recognise even his brother. Sacks asks all of his acquaintances to introduce themselves to him whenever they meet with him. He always requires an identity presupposition to see someone as x . But we are not like him! This might be a possibility, but not for us and not for all cases.

However, one could introduce a new notion of Identification-freedom for accommodating OI in addition to FI. On this suggestion, when we say that a judgement is identification-free, it should be free from both OI and FI. The question is whether this view is correct.

6.6. Object-identification does not always exclude IEM

Accepting object-identification may be harmless. It might be the case that 'identification' in identification-freedom means both object-identification and Frege-identification, and IEM excludes both. Identification-free judgment is free from both object and Frege-identification on that view. If this is true then all cases of object-identification have to open the possibility of misidentification. The question arises whether all object-identifications open the possibility of misidentification.

Unfortunately, there are judgments which are based on object-identification and which are still immune. Thus, not every object-identification opens the possibility of misidentification.

For a perceptual demonstrative I have to have an object which I perceive and to which I refer by using a demonstrative, like 'that one'. For using a perceptual demonstrative I have to perceive an object and take it to be the referent of that demonstrative. I cannot use the demonstrative for referring to

that object unless I think that the object falls under the individual demonstrative concept. For this reason object-identification is **required** for all uses of perceptual demonstratives. Object-identification is a necessary prerequisite to use a demonstrative.

On most accounts a demonstrative judgment like ‘That one is here’ or ‘That one looks terrible’ based on vision is immune relative to ‘that one’ (Evans 1982, Campbell 2002). But identification of the object (OI) plays a constitutive role in fixing the referent of the perceptual demonstrative. The reference of the demonstrative is fixed via attention and the subject’s intention to fix the referent of the perceptual demonstrative to the object perceived (Campbell 2002, 2007). Thus, nothing but the referent can be identified in that way for a perceptual demonstrative as the referent. Object-identification is the way to fix the referent of a demonstrative. So for a perceptual-demonstrative judgement, *Fd* (W), grounded on object-identification, if this OI fixes the referent of ‘d’ through W then such object-identification does not open the possibility of misidentification.

Why do we find a difference between Evans’s and Shoemaker’s notions of IEM? In part I of this chapter, I mentioned that Evans developed the characterisation of IEM for understanding the IEM of demonstrative thoughts and generalised it to understand the IEM of self-ascriptions. In contrast, Shoemaker worked out the characterisation of IEM primarily for immune self-ascriptions.⁷⁸ So for Evans – as we might expect – object-identification does not open the possibility of misidentification. This is because, for his paradigmatic case of IEM (the immune demonstrative thought), object-identification is necessary for fixing the referent. But for Shoemaker who is interested in self-ascriptions, object-identification is important; it opens the possibility of misidentification.

As I have shown object-identification does not **always** open the possibility of misidentification. For perceptual demonstratives, object-identification is involved in the reference fixing relation. It actually enables the subject to use a demonstrative term to an object. Thus ‘identification’ in the Identification-

⁷⁸ Evans’s discussion of immune self-ascriptions mentions identification-freedom only twice. He does not try to show how it relates to the IEM of self-ascription (Evans 1982: 221, 223).

freedom explanation of IEM cannot mean both object-identification and Frege-identification. Object-identification does not always open the possibility of misidentification.

One might suggest that every object-identification presupposes a Frege-identification. But this response on behalf of the Identification-freedom theorist would not work. There is a problem in thinking that object-identification requires or presupposes an identification. Suppose for the sake of the argument that object-identification presupposes Frege-identification. Object-identification is required for demonstrative thoughts. For the subject, to use a demonstrative requires thinking that a particular object falls under an individual demonstrative concept. Thus we have to have Frege-identification, if object-identification presupposes Frege-identification, for all demonstratives. Following this line of thought, if demonstrative thought can be immune, then for such thoughts Frege-identification does not open the possibility of misidentification. But this is worse for Evans, because then even Frege-identification does not always open the possibility of misidentification.

7. Summary for Part II

The identification-freedom explanation of IEM presumes that identification opens the possibility of misidentification. But there is an ambiguity in the notion of 'identification'. Identification-freedom does not seem to be the general explanation of IEM covering both first-person and demonstrative IEM.

The first reason is that what Shoemaker and Evans mean by 'identification' in identification-freedom is different. Shoemaker means object-identification and Frege-identification, while Evans means Frege-identification only. I have argued that object-identification is different from Frege-identification.

The second reason is that immune 'I'-judgments and 'that'-judgements are free from **different kinds of** identification. Identification-freedom cannot be a general explanation of IEM (for all kinds of IEM) if the notions of 'identification' are different depending on which kind of judgment ('I', 'that') is immune. But this is exactly what we found.

Third, it seems what explains the immunity of demonstrative thoughts ('that'-thoughts) does not explain the immunity of self-ascriptions and vice versa.

Consequently, I have provided three reasons to think that identification-freedom explanation cannot serve as a general explanation of immunity at all. Thus we need to find a general explanation for IEM. This is the task of the next chapter.

Appendix: Identification-dependent Immunity?

There are several dedicated ways of knowing like proprioception, introspection, interoceptive awareness (monitoring and registering the internal states of the body, like hunger), kinaesthesia, sense of balance, sense of pain, etc. Nothing excludes the possibility that I can use a demonstrative to the object of which I know through one of such way of knowing. If I use a demonstrative 'that one' to the object which I know through a dedicated way of knowing, then I might think 'that one is me'. This allows us to construct a case when my judgment based on a dedicated way of knowing is grounded on an identification (Frege-identification) but is still immune.

Identification-dependent Immunity

We can imagine a Buddhist monk who tries to get rid of the self. The monk only concentrates on sensations and tries to lose everything over and above the experience. (A neo-Humean might join in this exercise.) Suppose I am this monk. Consider a case in meditation when I know through proprioception that I have my legs crossed. I try to focus all my attention on feeling that; I am trying to grasp its phenomenology. I am forbidden to think about anything except my sensation. To develop such an exercise I will think of that object which I know through proprioception as 'that one'. I am immersed in my investigation and am writing down what is true of 'that one'. I do not pay attention to the fact that 'that one is me' for a split second. So I think that 'That one has her legs crossed'. I realize at once that I am that one, but my legs are not crossed any more. I think

‘I had my legs crossed’ based only on proprioception. (I am not a determined neo-Humean monk after all.) First, I think ‘that one has her legs crossed’ and second I think ‘I am that one’. Nobody except me can be the one whose legs were crossed. I knew it through proprioception a moment ago.⁷⁹ What I am trying to describe appears to be a possibility even if such a situation might never occur.

This example could be developed differently using actual cases of such enterprises. A possibility would be using a believer in introspection as a method of empirical study in psychology. Members of the fin-de-siècle Würzburg School involved in the ‘introspectionist movement’ in psychology would engage in such exercises of introspection, and it is quite possible that they concentrated only on the sensations and forgot the subject for a moment. This is possible because in some cases the knowledge I have that the pain is mine and that I know of myself may only be implicit.

I do not see what would exclude the possibility that my judgment is grounded on thinking about myself through two different dedicated ways of knowing and identifying of whom I am thinking about. Introspection and proprioception seem to give me the same object in different ways, but the object is necessarily the same. Thus I can identify the two objects by realizing their identity, but I cannot be mistaken in realizing such identity. This is the case even when I employ a demonstrative concept like ‘that one’. ‘I had my legs crossed’ is identification dependent because it is grounded on ‘that one is me’, and it is immune. If I am correct then I have found a case of identification-dependent immunity. I cannot misidentify who has her legs crossed, thus my thought ‘I had my legs crossed’ based on proprioception, introspection, and inference – including an identification component ($a=b$) – is immune. Here ‘Identification’ means Frege-identification.

⁷⁹ Let me set aside rewiring cases and assume that we are dealing with cases involving the kind of proprioception which we have and not a different, rewired kind. One might have her proprioception rewired to someone else and may be able to believe on this basis that that person’s legs are crossed. I discuss such cases in chapter 10.

On the face of it this case seems possible.⁸⁰ But it is excluded by any version of the Identification-freedom explanation of immunity. So this casts some doubt on the Identification-freedom explanation.

Evans requires that the 'Ideas' in $a=b$ have to be different. This means that the concept or the mode of presentation have to be different for 'a' and 'b'. The modes of presentation (MP) seem to be different in our example (a demonstrative MP for 'this one', a first person MP for 'me'). However, the thought 'that one^{known through proprioception} is me' is necessarily true. The modes of presentation seem to be different but the ways the referent of the two terms are fixed, even if they are different, necessarily provide *the same object* as referent. If this possibility is open, then identification-dependent immunity is possible.

One might think that it is **impossible not to be aware** that the object which I know of through proprioception is me. This is not so clear. Experiencing something involves attending to that thing. I can redirect my attention by will. I can attend to what my proprioceptive experience when I want to consider what is happening in me. The way that humans focus their attention allows that we screen out other aspects of the experience, so one might screen out the 'myiness' in the experience for a split second when focusing one's attention on the phenomenal qualities of one's experience. Phenomenological speaking there is nothing which excludes this possibility.

There are other reasons to support my claim. There are Frege-identifications which are necessarily true. For a square *the number of the sides* and *the number of the angles* are necessarily the same. The number of sides is a different MP from the number of angles, thus this is an example of a Frege-identification. It is possible in principle that for two referring expression, *a's* and *b's* reference is fixed differently but **their reference is necessarily the same**. It is possible, like in this case, that the subject could not doubt the truth

⁸⁰ One way to understand would be as follows. When I know of myself through proprioception the mode of presentation is a demonstrative mode of presentation: 'this object' or a hybrid: 'this object which is known through proprioception' (demonstrative-descriptive). I realise that I am this object based on introspection. Thus I have a mode of presentation through proprioception and a demonstrative. The subject takes it that they are representations of the same object and they are necessarily representations of the same object.

I try to show that if there is an identification in introspection then this identification would be harmless. The Identification-freedom theorist has to respond either that such identification is not possible or that such identification is not the kind of identification which is excluded by Identification-freedom.

of $a=b$ even though the MPs of the two terms are different. This means that here we have a case of Frege-identification where its truth cannot be questioned (with intact cognitive capacities). However, Frege-identification is supposed to be such that it is possible that the subject doubts it. But one cannot doubt that the number of the sides and angles of the square is identical. Similarly, in my example the truth of ‘that one^{known through proprioception} is me’ cannot be doubted without the subject having a cognitive deficit. ‘That one^{known through proprioception} is me’ seems to be a Frege-identification but its truth cannot be doubted without compromising the subject’s sanity. This is a Frege-identification which is necessarily true. Because ‘I had my legs crossed’ is based on a Frege-identification which is necessarily true and which cannot be doubted, I cannot misidentify whose legs were crossed on this basis. Consequently, I have found a case of identification-dependent IEM.

A defender of Identification-freedom will immediately reply with the following reformulation: identification-free thought, Fa based on W cannot be grounded on an identity ($a=b$) based on W except where $a=b$ is necessarily true. I do not object to this revision of Identification-freedom because it shows that what matters is not Identification-freedom, but how the reference is fixed. Moreover it directs our attention to the fact that knowing from the inside should figure into the explanation of the IEM of self-ascriptions. Thus, it seems that what bears the explanatory burden is the kind of way the subject knows of the object and its properties. I will develop this idea in the following chapters.

Chapter 7

The General Explanation of IEM and the Reference-fixing Relation

Abstract

I will argue that the general explanation of IEM has to employ the reference-fixing rule. To this end, I will show that, for a singular term 'a' in a judgement Fa (W), only the reference-fixing rule can rule out the possibility that nothing but a is what the subject thinks about as F (on basis W). This general explanation of IEM is able to accommodate the difference between the two kinds of IEM, demonstrative and first-person IEM. Their difference will be traced to a difference in the reference-fixing rules involved in each kind of IEM.

1. The Origin of the Reference-fixing Idea

In this chapter, I will argue that the reference-fixing rule should figure into the explanation of IEM. The idea that the reference-fixing rule should be a constitutive part of the explanation of IEM is not new. The reference-fixing rule figures into the general explanation of IEM for Peacocke (2008: 92) and Campbell (2013). But they suppose that there is no difference between the IEM of 'I'-thoughts and demonstrative thoughts. Peacocke, for example, uses the reference-fixing relation in his explanation of IEM, but this is assumed to be a general explanation of all kinds of IEM without allowing for a modal difference (Factual versus Strong Identification-freedom). The difference between my view and their views is the following. I argued that there are two kinds of IEM and I will offer an explanation of the difference between the two kinds of IEM, relying on the different ways in which the referent of 'I' and a singular demonstrative are fixed. On my proposal a judgement Fa (W) is immune iff the referent of 'a' is fixed through W (with further constraints). If this is correct then the reference-fixing rule figures into the explanation of IEM.

2. The Challenge of the Missing Unique Object

In discussing Evans, I distinguished the Special-basis explanation from the Factual Identification-freedom explanation. A version of the former was first introduced by Shoemaker (see chapters 4 and 6). According to the Special-basis explanation, for the subject, there are certain special ways of knowing which are private, and only a unique object, the subject, can be known through them. I called such ways of knowing **ways of knowing from the inside**. When the subject knows from the inside that she is in pain, hungry or out of balance, then the subject cannot misidentify who is in pain/hungry/out of balance. This seems to be a nice and simple explanation of the immunity of self-ascriptions. But this cannot be an explanation of the immunity of demonstrative thoughts based on multiple-object ways of knowing, as I argued in chapter 6.

This raises the following question: What can ensure, for an IEM demonstrative thought, $Fd(W)$, that nothing but d – only this unique object – can be thought to be F , on this basis W , if anything is F ? To explain the IEM of an immune demonstrative thought, $Fd(W)$, we have to explain why only a unique object and nothing else can be thought to be $F(W)$. This is easy for immune self-ascriptions. If, for Fi , knowing from the inside provides the subject only with a unique object, herself, then there is no problem explaining why, known only from the inside, nothing but the subject has to be F , if anything is F .

For demonstrative thoughts, the way of knowing is typically a multiple-object way of knowing. Through vision or touch, many different objects can be known. For demonstrative thoughts based on vision or touch, nothing ensures that only a unique object is known to be F . What would ensure that if I know that something is F through vision, or a multiple-object perceptual faculty, then there can only be one object which I can think is $F(W)$? This poses a powerful challenge for understanding how IEM of demonstrative thought is possible. Call this challenge for an account of demonstrative IEM, **the challenge of the missing unique object**.

This challenge divides into two questions:

1. How do we get the unique object for a demonstrative referring term in a demonstrative thought?
2. What could ensure for a perceptual demonstrative $Fd(W)$ that only the unique object d can be thought to be $F(W)$?

I will argue that the reference-fixing rule will answer the first question and how the reference-fixing rule figures into the general explanation of IEM will answer the second question.

3. Reference and the Reference-fixing Relation

Let me start by discussing the reference-fixing relation and how the reference-fixing rule for different kinds of terms differs. We use singular terms like proper names, demonstratives and the first-person pronoun; their function is to stand for objects in the world. This enables us to gain and store knowledge of objects in the world and revise our beliefs if they are wrong. A singular referring expression stands for a certain object and a singular thought contains such expressions. There are two questions related to reference which I will discuss shortly: (i) What does it mean to say that a singular term stands for an object? (ii) How is the referent of the singular term fixed to the object to which it refers? I now turn to the first question.

A singular term purports to refer to a particular object in a judgment, but not all singular terms succeed in referring. 'Vulcan' is one of such example. But most of the singular terms are used to refer to existing objects in the world; we call this subclass **referring singular terms**.⁸¹ Such a referring singular term succeeds in referring and standing for the referred object. What does it mean to stand for an object?

There is a referring singular term and an object which is its referent. The referring singular term makes this object available for the thinker to think about. We can think that the referring singular term is tagging the object. It makes communication about the object possible. This enables the subject, if

⁸¹ There are non-referring singular terms. When such a term is used, it at least purports to refer (e.g.: Phoenix, Vulcan).

certain further conditions are met, to acquire, store and integrate knowledge of the object and revise it.

Consider the following thoughts:

J1. Caligula was a roman emperor.

J2. Caligula was mad.

J3. This apple is green.

J4. I am skinny.

‘Caligula’ stands for an object, a certain person. J1 is true in case that person was a roman emperor and J2 is true if he was mad. When one uses ‘Caligula’, one uses it for that person, and may intend to refer to that person. In order to decide whether J1 is true we have to look at whether that person for whom ‘Caligula’ stands was a roman emperor. Different occurrences of ‘Caligula’ in different sentences in all counterfactual situations will stand for the same object (Kripke 1980).

Let me differentiate between the type and individual tokens of the first person pronoun. The first-person pronoun is a type, a token of which can be used on many occasions for different persons. Thus, when I use ‘I’, I use a token of it. When ‘I’ is uttered, in a particular situation, by a particular speaker, then what is uttered is a token of the type ‘I’. A token of an indexical refers to an object, place or time depending on its type (‘that’ – used for a perceptual demonstrative, ‘now’, ‘here’, ‘I’). The type ‘I’ can be uttered in different situations, but its referent is different if the speaker is different. Thus different occurrences of indexicals in different counterfactual situations can stand for different objects.

Similarly, different tokens of the same type of a perceptual demonstrative may stand for a different object. The token perceptual demonstrative: ‘this apple’ in J3 is uttered on a certain occasion; it stands for an object of which the speaker thinks about and intends to refer to in this situation. There is an object in experience which ‘this apple’ stands for. The token ‘I’ in the thought J4, in the situation when it is uttered, ‘I am skinny’, stands – as is commonly assumed – for the thinker/speaker. It is true iff the object which is me (if I am the thinker/speaker) is skinny at the time when this is uttered.

To sum up, a referring singular term, like 'Caligula', 'this apple', or 'I', serves in a language to pick out an individual object and can be used to think or communicate something about that particular object. A referring singular term stands for an object.

The question arises what a use of a referring singular term, a , involves. To use a referring expression, a , which stands for an object, \underline{o} , involves, roughly, the following conditions. (About my notation: \underline{o} is not a name for an object but it stands for the object itself.) The first is:

C1. The subject thinks about \underline{o} , the referent of a , in some way.

It is not sufficient that 'a' stands for an object. The subject has to use 'a' for an object in order to use it as a referring term. This condition requires that the subject knows which object she is thinking about in one or another way.

The second condition is this:

C2. For a thought Fa , the truth of Fa depends on whether the particular object \underline{o} which 'a' refers to is F .

We saw that the singular term 'a' stands for an object. For Fa to be true, that object – which the singular term 'a' stands for – has to be F .

A question arises how it is possible that a referring singular term stands for an object. Why is it not standing for something other than its referent or nothing at all?

For any use of the proper name, c , c will refer to the same thing.⁸² At least when we – in our linguistic community – do not know more than one person called Caligula, then 'Caligula' always refers to Caligula whenever it is used in our community. Suppose I forget Nero's name and use 'Caligula' instead of 'Nero'. Would it be the case that in J2 'Caligula' refers to Nero? Obviously not! So

⁸² Let us avoid difficulties arising from the fact that two people can have the same name (Burge 1973). Francis Bacon can be the philosopher or the painter and many other people can have the same name. Burge takes proper names to be indexical expressions. Proper names are types, so they have to be indexed and the context disambiguates which index we have to use for a token proper name.

J2 is true even if one utters it intending to say something about Nero. Consequently, there should be rules which fix their reference. Call these rules the **rules of reference-fixing**.

Similarly, I cannot use 'I' for something other than myself, and I cannot succeed in using 'here' for a place far away, or using 'now' for another time than the time of utterance. I might think today is my birthday so I can hope to get a cake now even though it will be tomorrow. I might try to use 'here' or 'now' for the wrong place or time, but I will not succeed in referring to the wrong place or time.

Demonstratives have a reference-fixing rule as well. On different occasions of use, a singular demonstrative referring term type can refer to different objects. In the same situation where there is an apple on the right and on the left, I might refer to the right or to the left apple by using a token demonstrative term ('that apple'). It seems that the subject's intention determines which object the referent is in a particular situation. The subject intends to refer to this object and what makes this possible is the way of knowing through which the subject knows of this object.⁸³

Thus we see that the reference-fixing rules for the different kinds of referring singular terms are dissimilar. I will rely on these differences in accounting for different kinds of IEM.

4. Toward a Solution

The first question of the challenge of the missing unique object was: How do we get the unique object for a demonstrative referring term in a demonstrative thought? I propose to answer the first question of the challenge of the missing unique object by considering how the reference of the perceptual demonstrative is fixed. Suppose a perceptual demonstrative judgement *Fd* based on vision is immune. For *Fd*, there is a unique object, the referent of 'd'. Only the reference-fixing rule for the perceptual demonstrative 'd' could

⁸³ One might deny this, but agree that there should be something which determines which object the referent is in the relevant circumstance (where the token was uttered). It is possible to hold that demonstratives refer to the most salient object in the situation where they are uttered. Salience can be restricted to such objects which are visible or audible from the subject's location. Even in that case, it cannot be arbitrary which object is the referent of a token demonstrative term.

provide a unique object which is necessarily *d* and cannot be anything else. What else could provide, necessarily, the object which is the referent of the demonstrative without the possibility of mistake?

A token demonstrative, 'd', is used in a particular situation. There are other objects in the situation where 'd' is used; thus something has to ensure that the object is *d* and not something else. The subject attends to an object in this situation and intends to refer to it by using 'd'. This fixes the referent of 'd'. But one could have used 'd' for another object. What ensures that the object which is *F* will be the object to which the referent of the demonstrative 'd' is fixed? Nothing, unless the reference-fixing rule is what determines the referent of 'd' in this situation. Only this can provide the unique object which is the referent of 'd' and cannot offer another object. The General Explanation of IEM of *Fx* will provide the answer to the challenge by exploiting the reference-fixing rule of 'x', allowing for the differences for different kinds of terms. But before showing that I will first look at how we get the unique object for demonstratives into our picture.

4.1. Where does the Unique Object Come From?

There are two options depending on the theory of how the reference of a (singular⁸⁴) perceptual demonstrative is fixed. Evans supposes that without having a unique object to which the referent of the demonstrative is fixed one cannot have a demonstrative thought.⁸⁵ Accordingly, for a demonstrative thought, if there is more than one or less than one object which one intends to refer to by using the demonstrative, then there is no demonstrative thought (Evans 1982: 46). In Evans's view, when there is a demonstrative term, then there is a unique referent. This is one possible way to provide a unique object.

We need not accept Evans's disjunctivism concerning demonstrative singular thought; I am neutral concerning this issue. I allow that one might hold

⁸⁴ I only discuss singular perceptual demonstratives; for convenience I do not mark this. IEM is a feature only of singular thoughts – as I discussed in chapters 3 and 4.

⁸⁵ Evans's slogan 'no referent - no thought' means that if there is no referent of the demonstrative term then there is no demonstrative thought. There might be another kind of thought but not a demonstrative thought. In a descriptive thought a singular term might not have a referent but for a demonstrative thought the demonstrative term has to refer to a unique object.

a different view. If the subject can purport to refer with a demonstrative without succeeding to refer, then IEM will only be a feature of the successfully referring demonstratives in demonstrative thoughts. This is another possible way to provide a unique object.

On all views, for singular thought with a referring term, whatever fixes the reference provides a unique object, the referent. But are there independent grounds for our claim that the reference-fixing rule matters for IEM? There is a case when it is clear that the reference-fixing rule plays a constitutive role in understanding IEM; I turn to this case now.

5. Julius – a Descriptive Name: ‘Julius’ refers to whoever invented the zip

I have discussed first-person and demonstrative IEM. But our list is short of one other kind of thought which can be immune. For Evans, descriptive names can be immune as well. In discussing Evans, I have yet to mention a contrast between descriptive names and ‘I’/‘that’.

The paradigmatic stipulation-based descriptive name is Evans’s example of ‘Julius’. Consider the following situation of introducing a new proper name into our language. Let us use ‘Julius’ for *whoever invented the zip*. We introduce this name to the language with this reference-fixing stipulation. We assume ‘Julius’ only refers if there was a unique inventor of the zip, but there is a thought even if it does not refer. Evans writes:

And even if we follow Russell, (...) there still remain terms which would intuitively be regarded as singular terms, but for which the ‘no referent – no thought (sense)’ position seems to be incorrect. A particularly clear example can be produced by introducing a name into the language by some such ‘reference-fixing’ stipulation as this:

Let us call whoever invented the zip ‘Julius’.

I call such names, whose reference is fixed by a description, ‘descriptive names’. (...) [I]t is not very plausible to deny that ‘Julius’ is a member of the same semantical category as other proper names. (Evans 1982: 31).

Notice that ‘Julius invented the zip’ is immune relative to ‘Julius’.

Let us have a closer look at what makes it the case that 'Julius invented the zip' is immune. What ensures its IEM is the way that the reference of 'Julius' is fixed. It cannot be the case that Julius did not invent the zip; Julius is just whoever invented the zip. This conforms to our original suggestion that the way that the reference is fixed plays a crucial role in explaining IEM.

However, descriptive names including stipulation-based names are atypical for the reason that they are very rare. The examples of descriptive names are 'Vulcan', 'Deep Throat', 'Neptune', 'Jack the Ripper' and 'Julius'. Proper names, demonstratives and other singular referring terms are not, in general, like that. In order to fix the referent of singular terms, we typically cannot find an appropriate description shared and known by the users of them (Evans 1982: 48-49).

There are certain reasons to hold that stipulation-based descriptive names are atypical for IEM. One might worry that even if descriptive or stipulation-based names in a judgement are immune, because their reference is fixed in an unusual way, this might be unique to them.

'Julius' as a descriptive name is atypical for immunity because it does not employ a way of knowing of the object which is the referent. Evans emphasises that there is no way of knowing involved in fixing the reference of 'Julius'. Evans distinguishes IEM in a narrow and in a broad sense. Thoughts about 'Julius' are immune only in a broad, but not in a narrow sense. IEM in the narrow sense requires that (i) the thought is identification-free, but, additionally, (ii) the thought has to be based on ways of gaining information from object(s) (Evans 1982: 31-32, 180-182). The cases of IEM in a narrow sense are different because they employ a way of knowing of the object.

5.1. An objection

An obvious worry here is whether I should generalize this result to other singular thoughts. Immune thought about Julius is not based on gaining information of an object in the world, but based on stipulation. So even if the explanation of IEM for descriptive names cannot be done without employing the reference-fixing relation, why should we have to use the reference-fixing relation to explain the IEM of 'I'-thoughts and demonstrative thoughts?

If we examine the special explanations of IEM in first-person IEM, demonstrative IEM and descriptive IEM, even though each explanation looks very different, there is one common aspect. The explanation in each case crucially alludes to how the reference of the relevant term is fixed. Stipulation-based descriptive names can figure in immune judgements because the reference is determined by stipulation (which involves description). In first-person IEM, immune judgements are possible because the subjects has special ways of knowing through which she knows of the referent of 'I' and through which she self-attributes qualities. In demonstrative IEM, the way a subject attends to an object in a specific situation and intends to refer to it by using a demonstrative fixes its referent.

Once we accept that the reference-fixing rule is critical to the explanation of IEM we can see that the differences between the different kinds of immune thought can be accounted for by the differences in their reference-fixing rules. Thus, such an explanation of IEM can not only account for the similarities between different kinds of IEM but also account for their differences. When the way of knowing plays a constitutive role in the reference fixing of the relevant term, then basis matters. For descriptive terms, ways of knowing do no figure in the reference fixing of the term, so basis does not matter. This explains why IEM in the narrow sense is basis-relative (chapters 1, 3, and 4) and that it is different from IEM in a broad sense. Consequently, involving the reference-fixing rule in the explanation of IEM accounts not only for the similarities but also the differences between Evans's narrow and broad IEM.

Thus we have good reason to believe that the reference-fixing rule should figure into the explanation of IEM.

6. Other Reasons for Using the Reference-fixing Rule in Explaining IEM

To provide further support for involving the reference-fixing rule in explaining IEM, I will examine three cases and show why not all of them are immune by using differences in the reference-fixing rule. This will demonstrate that the reference-fixing rule has to figure into the general explanation of IEM. Let me call these cases **Dog Affair I-III**.

In Dog Affair I, I see a dog that looks like it is barking (I do not hear it). Based only on vision I judge ‘That dog is barking’ (W =vision). My intention (or salience in my environment)⁸⁶ fixes the referent to the dog which I see and to which I attend. As we expect, based only on vision – through which its reference is fixed – I **cannot** misidentify which dog is barking. It is because the reference of ‘that dog’ is fixed by vision at that time and I only attend to that dog (or this is the only salient dog through vision at that time). Thus in Dog Affair I my judgment is immune relative to ‘that dog’ (W =vision).

In Dog Affair II, again I see a dog that looks like it is barking but additionally I hear the barking of a dog. On the basis of what I hear and what I see, I judge that that dog is barking (W =vision, audition). Vision provides us with the most detailed information and for evolutionary reasons our cognitive architecture is such that we mainly rely on vision if it is available to us in good viewing conditions (Burge 2010). So in good viewing conditions I mean the dog that I see by ‘that dog’; the referent of ‘that dog’ is fixed to the dog I see. Can I misidentify which dog is barking? Yes, indeed, I can misidentify which dog is barking. The audible dog might be barking while the visible one (which is a different one) is yawning. The reason why I can misidentify which dog is barking is this: the perceptual demonstrative’s reference is fixed to the dog which I see. Without invoking the reference-fixing rule we cannot explain what opens the possibility of misidentification in the second but not in the first case.

To demonstrate the force of this point, let me consider Dog Affair III. I only hear a dog barking. I assert ‘that dog is barking’ based only on audition. Could I misidentify which dog is barking? No, I cannot misidentify which dog is barking. In this case, the reference is fixed to the dog that I hear and misidentification cannot happen. There is only audition available to me to fix the referent to the dog.

My suggestion for what explains demonstrative IEM is this: when a demonstrative thought, Fd (W) is immune, it has to be based on that way of

⁸⁶ There are several objects in our environment but some of them are more salient. The most salient dog could be conceived as the dog which dispositionally attracts the most attention from a normal human viewer. Being the salient dog might fix the referent of a perceptual demonstrative without involving the intention of the user. I remain neutral but find the intention-involving story more plausible.

knowing through which the reference of the demonstrative, *d*, is fixed (when it is fixed). Relying on this, I can explain what opens the possibility of misidentification in Dog Affair II (the case which is not immune), but not in the other two affairs which are immune cases. In Dog Affair II the judgment is based on audition and vision. More precisely, the way of knowing on which the attribution of barking is based involves audition while the way of knowing through which the referent of the demonstrative is fixed is vision. This opens the possibility of misidentification.

In Dog Affairs I and III, the judgments are immune because the basis (way of knowing) is restricted to that basis through which the referent of the demonstrative is fixed (at the time when it is fixed). This is not the case in Dog Affair II. What opens the possibility of misidentification is the basis of the judgement not being identical with that basis through which the referent is fixed. This is what we would expect, but only if the reference-fixing rule is constitutive in the explanation of IEM.

One might object that we should hold that 'that dog' in Dog Affair II is fixed both by vision and audition. But even in this case there is no item which can be immune because the reference is not fixed to a unique object. (Either the demonstrative only purports to refer or, on Evans view, there is no demonstrative thought. But for IEM we need a thought employing a singular *referring* term relative to which it can be immune.) But again the explanation should and will rely on how the reference of the perceptual demonstrative is fixed. Involving the reference-fixing rule is a natural solution. This is exactly what I suggest.

7. The solution of the Missing Unique Object and the General Explanation of Immunity to Error through Misidentification

The missing unique object is a challenge for explaining how it is possible that for an immune demonstrative judgement Fd (*W*) only a unique object can be thought to be *F* on this basis *B*. For an immune demonstrative judgement Fd (*W*), there is a unique object the referent of '*d*'. But it is not yet clear what can ensure that only *d* can be thought to be *F* on this basis *W*.

Nothing but a can be F iff for Fa (W) the reference of ' a ' is fixed through W and its F -ness is known through W when its reference is fixed. This explains why we have a unique object the referent of ' a ' and why nothing but a can be thought to be F (W).

Accordingly, my suggestion is that, for Fa , when one gains knowledge of the object, a , and of its F -ness through a way of knowing W_1 at time t_0 (where W_1 is the way through which a 's reference is fixed) then one cannot misidentify which object is F . If the only way to know of a and its F -ness is through W_1 , at the time when its reference is fixed through W_1 , then nothing but a can be known to be F through W_1 at that time. This is the general explanation of IEM.

How can we account for the difference between first person IEM and demonstrative IEM? A special explanation is an explanation of the immunity of one kind of IEM, e.g. first person IEM. In contrast a general explanation is an explanation for all kinds of immune thought. I turn to discuss the special explanations of demonstrative and first-person IEM.

In chapter 6 we saw that the time of the reference-fixing and attention to a unique object matters for demonstrative thoughts but it does not matter for self-ascriptions. The reason for this is that the subject knows of herself from the inside. Only a unique object, the subject can be known from the inside, so one need not keep track and discriminate this object from other objects (Peacocke 2013). As I have shown, the difference in conditions of IEM reflect the conditions of how the reference is fixed in each case.

Using a demonstrative, d , for an object which one sees and attends to, and judging that it is F , is sufficient to make Fd (W =vision) immune relative to ' d '. When one visually attends to an object and uses a demonstrative for it, then the way the reference of the demonstrative is fixed provides the unique object which one can think is F .

The reason we have immune thoughts, for demonstrative thoughts and self-ascriptions, Fa (W_1), is, I suggest, that a 's reference is fixed through W_1 .⁸⁷ I have

⁸⁷ I do not rule out that Evans might have something similar in mind and this explains why the Special-basis explanation is different from Factual Identification-freedom. Although this cannot be unless Evans would accept that first-person and demonstrative IEM is of different kind. There are some reasons to think that Evans used IEM to find how the referent of a referring term is fixed. Evans says that the referent of 'that', 'here' and 'I' are not fixed by a definite

argued the explanation of IEM of an 'a'-thought requires the reference-fixing rule for 'a' as its constitutive part.⁸⁸

description in contrast to 'Julius' which is fixed by a description. Evans's difference between narrow and broad IEM can be viewed as supporting this claim.

Evans used the notion of 'fundamental identification' for fixing the reference. Evans only mentions that the notion of fundamental identification is different from the notion of identification involved in Identification-freedom. He never says that the fundamental identification should figure in the explanation of IEM.

⁸⁸ This does not cover the explanation of the IEM of descriptive thoughts, but these cases are less interesting. The real general explanation would be Fx is immune iff the way as the reference of x is fixed (which may or may not involve a genuine way of knowing) excludes that something else than x is F .

Part II

The Simple View

I am now in a position to propose an answer to the question of how the reference of 'I' is fixed. My proposal will be called the *Simple View*. There will be three parts to the exposition of the view in three chapters. In chapter 8, I introduce the Simple View. In chapter 9, I will use a test to argue for the proposed view. IEM of self-ascriptions will be our guide in the test of the explanatory power of the Simple View. I will try to argue that unless the referent of 'I' is fixed as I propose, IEM cannot be explained. To set up the test, I will list several essential features of uses of 'I' which should be accounted for on all views about how the reference of 'I' is fixed, including GRR, IEM and aboutness-error freedom. I will demonstrate that the Simple View can explain all the listed essential features of the uses of the first-person pronoun, but other views cannot. A view is a good candidate to be the correct one only if its defender provides answers to the basic objections against it. In the final chapter, I answer the most pressing objections to the Simple View.

Chapter 8

The Simple View: Fixing the Reference of 'I'

Abstract

How is the guarantee against incorrect reference possible? According to the guarantee against incorrect reference (included in GRR) the subject, by using 'I', cannot intend to refer to the wrong object; thus she is guaranteed to refer to the correct object. If uses of 'I' are required to have guaranteed right reference, then the subject has to be presented with the object that she is without the possibility of mistake. How is this possible? It requires that there should be some ways of knowing which enable the subject to be directly acquainted with herself without the possibility of mistake about which object she is. 'I' is used to refer to the object which the subject is directly acquainted with from the inside.

1. The Building Blocks of the Simple View

How is the guarantee against incorrect reference possible? I want to uncover what enables GRR. To do this, let me begin with some observations about the character of the object that 'I' refers to – but which describe the object independently of its being the referent of 'I'.

1.1. The Object Under Permanent Care: the Exposed Object

According to the Simple View, the referent of 'I' is the unique object that the subject knows about through special private ways of knowing. However, that object should be characterised independently of being the referent of 'I' and from being the subject itself. Simply saying that for any use of 'I' the referent of 'I' is what 'I' refers to or the subject is the referent of 'I' would not be very explanatory. If we can characterise the referent of 'I' independently of the first-person point of view, then we can check whether the referent of 'I' is fixed to the right object, according to my proposed view.

Let me look at an organism with cognitive abilities. Such an organism is typically able to take care of itself when it is awake. It is able to avoid danger and fulfil its desires in a more sophisticated and flexible manner, which is decoupled from environmental affordances (e.g. food affords eating), as compared to an organism without cognitive capacities. We can say that such an organism is able to take care of itself. To 'take care of x' at least means to "sustain or increase x's potential to survive".

For every organism, although it can care about other organisms, there is a unique object which it has to constantly care for. In order for an organism to be able to take care of any other organism, first of all, it has to permanently take care of itself. We are not surprised that the object requiring permanent care is the organism itself. The organism itself is the **object under permanent care**.

For an organism, the ability to take care of itself requires that it knows of the object under permanent care. This ability requires that an organism have a relation to itself which is manifest in certain behavioural patterns, such as self-directed actions. To perform self-directed actions like eating, drinking or scratching, the organism has to know of the object under permanent care.

The object under permanent care is the first building block of the Simple View.

In psychology there is a very useful notion explaining how an organism can survive: *having drives*. A drive is something governing certain kinds of behaviour (including actions) of an organism in the relevant situation. Such kinds of actions are usually necessary for survival and reproduction. The drives are dispositions to act in suitable situations for sustaining the organism's life preserving functions. Although the behaviour of the organism is driven by these drives, drives do not require that the organism knows of itself (as itself or *de se*). There are plenty of drives to take care of an organism; feeling thirst or hunger will drive the organism to drink or eat in suitable situations.

The object under permanent care is described relative to an acting animal. For an acting animal, the animal itself will be the object under permanent care. Animals collect food, eat, drink and reproduce. We might think of drives as a collection of capacities enabling these sorts of life sustaining actions.

For humans, the object under permanent care will be the object that is the referent of 'I'. A human subject can care about this object even if it is not yet able to employ the first person pronoun in thinking of itself.

A crucial step is when the subject does not merely know of the object under permanent care in a way that is manifest in self-directed action, but when it is able to consider the object under permanent care *as an object*. One marker for the emergence of this capacity is when an organism passes the mirror test (compare Peacocke 2014, where he discusses 'level one' *de se* content). In the mirror test, the experimenter puts a red dot (for children and apes) or a yellow line (for elephants) on their face. They are able to see this in a mirror. Passing the test requires that they remove the mark from themselves. This behaviour is taken to demonstrate that these creatures have implicit *de re* knowledge of themselves. The subject's use of the first person pronoun might be anchored to this preliminary existing implicit *de re* knowledge of the object under permanent care. (This is suggested by the developmental trajectory of children. They pass the mirror test before they are able to use 'I'.) The object under permanent care is the object which is the subject (in normal conditions).⁸⁹

It is necessary for the subject to be in a position to always care for that object (under permanent care) so as to survive, therefore a subject has to permanently know of that object. Introspection, proprioception, kinaesthesia and several other ways of knowing enable the subject to know of that object. But internal ways of knowing are not the only ways a subject can be available to herself. Vision, audition and touch provides information of the subject to the subject most of the time when the subject is awake. In a sense, that object under permanent care is permanently exposed to the subject.⁹⁰ When the subject is able to know of the object under permanent care as an object, I will call the object under permanent care **the exposed-object**.

⁸⁹ Fainting or being in a coma, and similar conditions do not count as normal conditions. In such conditions the subject is not able to survive alone.

⁹⁰ The subject is so much exposed to herself that there is a mechanism which reduces, for the subject, the available information of herself except in danger or situations when it is necessary to monitor whether things are as they are expected. This mechanism is called sensory attenuation (attenuation of information concerning the subject). For further discussion look at Frith (2005).

1.2. Cassam's Tool: Ways of Knowing

In order to state the Simple View, I employ Cassam's (2007) notion of *ways of knowing*. Ways of knowing enable the subject to acquire knowledge or beliefs.⁹¹ Let me go through a list of different ways of knowing and describe what their function is before I turn to discuss the dedicated ways of knowing. The subject has several ways of knowing of objects, events and state of affairs. Examples of ways of knowing are vision, audition, touch, testimony, proprioception, kinaesthesia, inference, memory and introspection. These ways of knowing are ways through which one can gain knowledge of objects and their properties or events.

To get the relevant way of knowing we might just ask the subject when she knows p 'how do you know that p ?'⁹² She might answer that she knows because she sees it (vision) or she remembers it (memory) or she had heard it somewhere (testimony). Whenever one believes that p but there is no answer whatsoever to the question how do you know that p , then p cannot be knowledge. In order to know that p there should be a way for the subject through which this knowledge is acquired. It might be the case that the subject does not know in which way exactly she knows that p . However, for Cassam, if the subject knows p , then there has to be a way of knowing that provides the answer to the question 'How do you know that p ?' even if the subject is not aware of it. (Note that my claims are restricted to values of p which concern a singular thought of a particular object. Some general thoughts might be acquired *without* a particular basis – way of knowing.⁹³)

⁹¹ Shoemaker (1968: 547, footnote 7) and Evans (1982: 180) both use the notion of ways of knowing in their discussions of IEM, but only very occasionally.

⁹² It is a question whether a subject can be completely ignorant of which way one gains knowledge. Still, it might be that only an ideal subject can always answer the question 'How do you know it?' but the average subject does not always know this answer. But I take it that we typically know the answer.

⁹³ One might object that there are several things which we simply know like 'Elephants don't wear galoshes in the wild' without a way of knowing dedicated to it. The reason why we know this about elephants is that we infer from things which we know that this should be the case. Moreover, we acquire most of our general thoughts through inference. When I discuss a way of knowing I am thinking mainly about singular thoughts, particular events and individual objects.

A belief like 'the world around us and objects in it exist' might be such that we were born prepared to believe it. There seem to be no ways of knowing for this. Self-identity could be another similar example. Such general thoughts seem to be presuppositions of most of our other thoughts. We presuppose them in some way. Thus I do not hold that there cannot be

It is not the case that any answer should be accepted to ‘how do you know that p ?’ There are certain methods like guessing or wishful thinking which might guide action or thought, if the subject wishes to rely on them. But guessing never provides knowledge to the subject about something other than ‘what would I guess now?’ The question ‘How do you know that p ?’ cannot be answered by saying I guessed it or I know it through wishful thinking. What I guess or I am wishfully thinking might be true. But knowledge cannot be a happy accident. These are not such epistemologically useful relations to objects or state of affairs, which would make knowledge possible. Accordingly, guessing and wishful thinking are not ways of knowing of objects in the world.

This thought is similar to what Evans was after when he required channels of information for knowledge of objects (except for stipulative names).⁹⁴ But I find the notion of ‘way of knowing’ more precise because my brain has channels of information which could never provide me directly with conscious content (e.g. the monitoring of blood pressure). The brain monitors changes but not all channels of information involved in this enable the subject to know of objects (including herself).

Summing up, there are two contrasts here. First of all, ways of knowing are such that they enable knowledge. Second, knowledge is not merely information that is available to the subject’s cognitive system in some way.

Now I turn to the second point, the difference between information and knowledge. Ways of knowing are those channels of information which can provide conscious content to the subject. There are channels of information providing only content which is never available to the subject. A channel of information which provides knowledge is a way of knowing. But there are channels of information which are not ways of knowing.

A subject’s heartbeat rate is monitored by the subject’s brain, although, if you ask a subject to report it, you might find people who do not have a good sense of their heartbeat rate. So it seems in such cases the subject relies on guessing. Thus this information is not available through a dedicated way of knowing for this subject. Even so, the subject’s brain relies on this information

knowledge which is acquired without employing ways of knowing, even though these cases are exceptional.

⁹⁴ Recall descriptive names like ‘Julius’.

and regulates heartbeat changes if it is necessary. In this case, what is computed over in the brain is information and not knowledge or belief.

What enables the subject to know of herself could be a private way of knowing of herself – as we learned from Shoemaker in chapter 4. Shoemaker writes: “[E]ach person knows of his own mental or psychological states in a way in which no other person could know of them” (Shoemaker 1968: 565). I follow this suggestion and try to find those ways of knowing of the subject which are ways of knowing only of the subject. What are the constraints which make the subject’s directly acquaintance with the exposed object (herself) possible?

2. Constraints on those Ways of Knowing which would Provide Direct Acquaintance with the Exposed Object

I turn to discuss the constraints which provide us with dedicated ways of knowing of the subject. My guide will be the guarantee against incorrect reference. My guiding question will be: **how is it possible that the subject knows of an object (which is herself) without the possibility of mistake?**

2.1. The First Constraint: the Privacy Constraint

Frege claims that there is a difference concerning how I know of the object which I am and how I know of other objects in the world. There is a privileged way of knowing of the object that is the subject which is only available to the subject. “Now everyone is presented to himself in a particular and primitive way, in which he is presented to no-one else” (Frege 1956: 17).⁹⁵ I called such a way through which the subject is acquainted with the object, where it is not accessible to others, **a private way of knowing**. Shoemaker puts this point like that: “Now there is a perfectly good sense in which my self is accessible to me in a way in which it is not to others. There are predicates which I apply to others, and which others apply to me, on the basis of observations of behaviour, but which I do not ascribe to myself on this basis, and these predicates are precisely

⁹⁵ I do not discuss the interpretation on which ‘being presented’ means a peculiar first person mode of presentation.

those the self-ascription of which is immune to error through misidentification” (Shoemaker 1968: 562).⁹⁶

According to privacy, there is a privileged way of being presented with the object which is myself. I might add that there is more than one private way of knowing.⁹⁷ As I put it, there are ways of knowing of the object, myself, which are not available to anyone but me. Others cannot know *of the object, me*, through these ways.

Let me elaborate the idea a bit further. Suppose there is an object *o* which I see in good lighting conditions. Vision, audition, touch are such senses where anyone, in the same spatiotemporal position as I am, equipped with reliably working senses, would similarly be able to perceive and thereby be able to know of any object which I perceive. In contrast, introspection, kinaesthetic sense, proprioception and nociception are not at all like vision. If I feel a stomach ache through nociception, it is not sufficient for others to be in the same spatiotemporal position as me to feel it.

Only the subject can know of the exposed object and its features through these ways of knowing. Note that I am merely recording a fact about our nature when I observe that there are private ways of knowing.

2.1.2. The Privacy Constraint

What does a private way of knowing consist in? Privacy restricts the ways of knowing to those through which only the subject can know of a particular content, *C*, through that way of knowing, *W*₁. This does not exclude the possibility that others can know of that particular content, *C*, through another way of knowing, *W*₂.

⁹⁶ I called this the special basis explanation of IEM. This can be read as requiring such self-knowledge to be non-perceptual. A judgment would be immune even when based solely on proprioception or kinesthesia (perceptual faculties). Thus the non-perceptual interpretation is wrong.

⁹⁷ For Shoemaker, private ways of knowing are non-observational and non-perceptual, but I do not suppose this. Martin (1993) argued that pain is perceptual faculty and Burge (2010) has argued, based on the lack of constancies, that pain and proprioception are non-perceptual. Constancy is required if changes in lighting or other conditions require it so that we can track an object as the same across different conditions. With this in mind, I do not see why something like that would be required for pain to fulfil its function. I treat nociception as a perceptual faculty but this is not necessary for my arguments.

In my reconstruction, Frege's suggestion is that privacy might be a natural constraint on a way of knowing which only presents the subject to herself. So let us take privacy as the first constraint on an appropriate way of knowing. But one might certainly entertain some worries about privacy; now I turn to discuss them.

2.1.3. A Worry about Privacy

One might worry that some brain scanning technique might turn a private way of knowing to a public way of knowing. By scanning the brain it seems that anybody can access any content which is in the brain, including information available to the subject by private ways of knowing. This is right. Through brain scanning the content will be available to anybody in an appropriate position (viewing the scanner), but it will be not available **through the same kind of way** as it is to the subject. Suppose I am aware of my pain by feeling it. It is not the content that is private, and it is not the intentional object that is private. The access to that content is private and this is a crucial difference. Frege's thesis, as I understand it, is about the access. The content like 'I have a stomachache' is available by report, so even testimony can deliver the content gained through private way of knowing to others.

2.1.4. Privacy is not Sufficient

My question is how the subject can know of the object, the referent of 'I' without the possibility of mistake. We have to test the idea that private ways of knowing provide the subject for the subject in the required manner. If privacy is sufficient then **the only object** which I could know through a private way of knowing **should be me**.

Let us suppose I am pregnant and near term with a baby. I know of my baby and its movements and hiccups by experiencing strange things from the inside. He likes to make somersaults. I know exactly when he makes a somersault, and

nobody can know this in the same way as I know it. I know of his somersault because I know of him, but still I know it from the inside.⁹⁸

Privacy requires that only the subject can know of a particular content (e.g. baby is somersaulting), C, while others cannot know of that particular content, C, through that way of knowing, W₁. Privacy is satisfied for knowledge of the baby and his movements. (One can know that she has a baby without testimony, but only through feeling it from the inside and from general facts.) I can know of my baby and its movements from the inside. But I am not my baby. Thus it is not the case that privacy is sufficient to provide only the exposed object without the possibility of mistake, as GRR requires.

One may worry that I know of my baby through inference and inference is not a private way of knowing. Even if I know it through inference, this inference is only based on a private way of knowing – if I infer that I have a baby based on what I feel from the inside. In this case even if this knowledge were based on inference, its primary source would still be the private way of knowing. If privacy were sufficient for knowing of the exposed object then privacy plus inference should be sufficient too. But this is not the case. Thus privacy is not sufficient to ensure that I know of the exposed object.

One might object that the relevant experience is a tactile experience and touch is a multiple-object way of knowing. This might be right; however I need not deny this in general. Touch on occasion can be private way of knowing. The only claim here is that the way of knowing relative to these contents (knowing of the baby and her movements) are private and not available to others. Consequently, privacy on a way of knowing is not sufficient to ensure a unique object – as the guarantee against incorrect reference requires.

If the subject knew only of a unique object through some private ways of knowing this would seem to be sufficient for a guarantee against incorrect reference. When we discussed Shoemaker and Evans, we found that there are some ways of knowing through which only a unique object can be known and Martin offers something on the same lines. This will be the second constraint: the unicity constraint.

⁹⁸ Ma case is inspired by O'Brien's internal baby-monitor. This device only provides information for the mother of the baby in the womb (O'Brien 2007). The internal baby-monitor has privacy and unicity (the next feature) so I will discuss it later.

2.2. Martin's Point: the Sole-object View

2.2.1. The Second Constraint: the Unicity Constraint on Way of Knowing

M. G. F. Martin (1995) put forward a very influential proposal according to which bodily awareness is awareness of a unique object: the living body. Proprioception, kinaesthesia, sense of heat/balance/pressure/pain and many other channels of information are integrated into bodily awareness. Bodily awareness cannot provide anything but information of a unique object and its parts.

One criterion for whether some form of sensory awareness constitutes a perceptual faculty is whether illusion is possible. Bodily awareness is subject to illusion. In the case of the phantom limb the subject thinks there is a limb which is part of her body even though the limb has been amputated. After surgery and in war this is a very common illusion or delusion. The possibility of illusion provides a good reason to believe that bodily awareness is a perceptual faculty.⁹⁹ According to Martin, bodily awareness is a perceptual faculty and only one object and its parts can be known through it.

How do we get the sole object according to Martin? On Martin's view the sole object has to be determined in a mind-independent way; it cannot just be whatever I am aware of through bodily awareness. Otherwise the thesis would be uninformative.

My solution is to use the notion of 'exposed object'. This is characterised independently of being experienced from the inside or being the referent of 'I'.

2.2.2. The Unicity Constraint

The unicity constraint restricts ways of knowing to those through which only one object can be known. I count nociception, sense of balance, introspection or proprioception as different ways of knowing. Unicity, as I construe it, will be a constraint on ways of knowing.

⁹⁹ I think bodily awareness is perceptual but I do not assume it. If it were not perceptual it would be still true that only that object which is the subject can be known through it.

2.2.3. A Worry

A pressing worry that I have to face is *the many parts worry*. I am aware of many parts of my body so it seems that it is not a single object which I am aware of any more.

In reply: First, note that all my body parts are only derivatively identifiable. There is a representation of the body: the body model (Tsakiris 2010, Longo and Haggard 2012, Wong 2015). The parts of the body are identified relative to the body model. So it is not the case that there is more than one object that I can be aware of. But my attention to the very same object can be different. I might pay attention to the whole object or I might pay attention to its parts.

The existence of the part is in some sense dependent on the existence of the whole object. Roughly, part of *x as part* cannot exist unless *x* exists, thus part experienced *as part* is ontologically dependent on the whole. Parts cannot be understood unless the whole is in some sense understood; thus knowledge of parts is epistemologically dependent on presupposition of the whole.

Second, there is a difference in the experience of body parts when experiencing them as parts or as a whole. We can see this when we consider the experiences of body ownership in two different pathologies: somatoparaphrenia and alien limb syndrome. Patients suffering from somatoparaphrenia do not feel a sense of ownership over their hand. They deny that one of their hands is their own hand, but assume that it has an owner (Vallar and Ronchi 2009). They believe that it is someone else's hand (e.g. husband/doctor/niece). Thus they experience the hand as part of someone else – not as a whole.

There is another delusion showing that one might experience her hand as a whole but not as a part. On the other hand, in alien limb symptom, patients do not feel a sense of ownership over their hand, but do not attribute it to themselves or anyone else (Vallar and Ronchi 2009; de Vignemont 2007, 2011). They experience an object, namely, an alien limb, attached to them. Such a patient is aware of her hand as a whole, a limb without an owner, but not as a part. Thus even in experience there is a difference between being aware of a part as a part or being aware of a part as a whole.

When one experiences her hand as her hand then she is typically aware of herself from the inside. When one experiences a part as a part of a whole then one thereby experiences the whole. Thus it seems that the only object which is available to the subject through private and unique-object ways of knowing is the exposed object even when one only experiences parts of it.

2.3. Privacy and Unicity are not Sufficient

Are the two constraints privacy and unicity sufficient for providing the exposed object (the subject)? Let us consider the following situation for making it clear why privacy and unicity are not sufficient to get the exposed object.

There is interesting research where scientists build brain-computer interfaces.¹⁰⁰ In such cases there is a silicon chip built into the brain which can read what one thinks and which can be used as a source of information, an artificial way of knowing. After a short training phase one can learn to gain information through this device and it is possible to control something built into this closed system through this chip – even a helicopter. The example is based in reality; brain-computer interfaces were designed by the Pentagon to enable pilots to have faster reaction times when they fly a helicopter or an airplane. This would result in a reduction of error. An apparatus reading brain waves translates the patterns of firing in the brain to pre-set actions which the computer implements.

Suppose I have a computer chip in my brain which is of a unique design – it only works with my brain, and I can monitor a super advanced helicopter's state with this gadget. I can fly the helicopter as well. Let me call this example **Helicopter from the Inside**. I have a private way of knowing of the helicopter that cannot be available to anybody except me in this way. I know only of the helicopter through this artificial way of knowing. It is not a natural design. But it is a way of knowing which enables me to gain knowledge of this helicopter.

¹⁰⁰ Another example is the internal baby-monitor which I mentioned earlier (O'Brien 2007). This device has privacy and unicity but does not provide the subject for the subject. My argument uses the same strategy which was developed by O'Brien to argue against a certain characterisation of knowing from the inside. Privacy and unicity are not sufficient to restrict the ways of knowing to those which only provide the subject for the subject.

In this case, I know of the helicopter through a private unique- object way of knowing. Does this make me the helicopter? Obviously, the helicopter cannot be me. Consequently, knowing an object through a private and unique-object ways of knowing is not sufficient to ensure that what is known is only the subject and nothing else – as guarantee against incorrect reference requires. Privacy and unicity as constraints on ways of knowing do not guarantee that the only object which could be known is the exposed object.

2.4. The Third Constraint: Internality

What ensures that through proprioception, kinaesthesia, or sense of balance only the subject can be known? If they are internal channels of information, then, for the subject, only the subject can be known through them. Let me call all systems which store, use and update information **computational systems**. For a computational system, S, a channel of information which only monitors S is **an internal channel**. For a computational system S, when S has a channel of information C which only monitors states of affair (processes and events) in S and only S, then C is **an internal channel of information**.

When my computer flashes, indicating that it is out of battery, or a car shows that it is out of oil, then this information is coming from an internal channel of information. Some of the channels of information in humans (a computational system as well) are internal channels of information. When a group has a closed-circuit television or phone then these are internal channels of information relative to that group. What makes it the case that proprioception, kinaesthesia, introspection or sense of balance and heat (and several other special ways of knowing) cannot be of anything but the subject is this: they are internal channels.

Internality is a constraint on channels of information and not only on ways of knowing. A channel of information might not be a way of knowing so in this case it cannot provide the subject for the subject. For example, channels of information provide heartbeat, blood pressure and sugar level monitoring, but these channels of information never become ways of knowing in their own right for the subject because they do not provide conscious content.

2.4.1. A worry

But is it not the case that internality as a constraint on channels of information alone is sufficient to provide the exposed object as the unique object which can be known through this way? No. First of all through an internal channel of information many objects can be known, like through a closed-circuit camera or phone. A group of doctors or a rescue squad might operate in a noisy environment and use a closed-circuit phone to communicate only with each other. The channel of information is internal, but all the doctors/rescuers (objects) and what they say can be known through it. Internal channels of information neither need to be private nor unique object channels.

Internality does not require that there is a subject on the receiving end of the information. Internality requires unicity to ensure that it is information about a unique object and privacy for ensuring that it is information for the subject. Internality ensures that the one *of whom* the information is about and the one *for whom* the information is for cannot be different. Thus internality for channels of information of the subject is not sufficient for providing knowledge of a unique object.

2.5. What is Missing?

Are we done? The question is whether these constraints – privacy, unicity and internality – are sufficient to put the subject into the position to know that *she knows of herself* through such special ways of knowing. For using the first-person pronoun we find it plausible, as Evans (1982), Peacocke (2008) and O'Brien (2007) do, that the subject knowingly and intentionally refers to herself. The subject knows that she is referring to herself and the subject intends to refer to herself by using 'I'. But can the subject know that she knows of herself whenever she knows of an object through private (P), unique object (U) and internal (I) ways of knowing?

Let me call such ways of knowing which are private, unique object and internal ways of knowing: **self-reflexive ways of knowing**. They ensure that only the exposed object (the subject) can be known through them. Accordingly, self-reflexive ways of knowing are necessarily of the exposed object, the subject. However, being self-reflexive for a way of knowing does not seem to ensure that

the subject knows that it is of herself. We need more in order to put the subject in a position to be aware that the content – provided by a self-reflexive ways of knowing – is of herself.

2.6. The Last Constraint: a Constraint on Awareness

There should be some kind of knowledge which enables the subject to self-ascribe what is known through self-reflexive ways of knowing. Which kind of knowledge would be able to govern self-ascription?

Consider the following story. Suppose an old friend of yours, Adorno, approaches you on the street in a foggy day and whispers into your ear ‘I feel your pain’. Adorno is an economist; he means what he says. Adorno means it non-metaphorically as always. He thinks what he feels is your pain. You might wonder whether he puts his non-existent poetic hat on, but you should exclude this firmly. So, you would be pretty worried. You would prefer to have a psychiatrist at hand, or seek help from a doctor or at least not be alone in the dark with an economist who seems to feel your pain. This is surprising! We *do* make mistakes and we are not referred to the psychiatrist for our mistakes. Normally, we are not distressed by mistakes or false beliefs of others.¹⁰¹

Compare this case with another one. Suppose on another foggy day that Adorno goes to you and whispers to you that ‘There are no black swans’. Nothing would make you worry! You will be neither surprised nor disturbed. You would not wish that a psychiatrist should be around and help. You may send Adorno to Regents Park to look at the two black swans on the lake far from the city. Similarly, if Liz says ‘Marble Arch does not exist’ you just point to the Marble Arch to change Liz’s view on that matter. You have no similar move to make to show that one cannot feel (non-metaphorically) the pain of others. Mistakes can be corrected but not mistakes of the first kind. Thus you cannot direct Adorno’s attention to the evidence or ground which is only available to him.

¹⁰¹ There are some exceptions but they are of different kind. If one would say ‘I figured out that $2+2=5$ ’, you would be pretty worried as well. But this is failure of basic rational capacities and this makes it disturbing. This is a less likely (but possible) explanation for the economist friend case as well. Moreover nobody ever dared to think that ‘I have pain’ is analytic while one might hold this for ‘ $2+2=4$ ’.

The evidence (reason or whatever grounds or justifies it) seems to come from the inside and it cannot come from anywhere else.¹⁰² Let me elaborate what I mean by evidence. If I say 'I have a headache', but I do not feel anything then I am not justified/entitled to self-ascribe a headache. I call 'evidence' whatever is the difference between the case when I feel a headache and the case when I do not feel headache. In the second case I lack evidence for the self-ascription of headache.

It seems that for normal people, when one knows of a state from the inside through a self-reflexive way of knowing then one, thereby, is in a position to self-ascribe it. Something makes the self-ascription automatic. Such self-attribution cannot be defeated by evidence for attributing that state to someone else. I might not have pain when I seem to feel it, but it cannot be the case that what I feel from the inside is someone else's pain and, as a mistake, I misattribute the pain to myself. Let me call what makes this possible, which is a constraint on awareness: **automaticity**. Unless one is delusional or rationally impaired, automaticity *enables* self-ascription of states known through self-reflexive ways of knowing. Automaticity *enables* the content known through self-reflexive ways of knowing to be self-ascribed, given that the subject is able to use 'I'. But automaticity (registration of self-significance) is already present before one acquires the mastery of 'I'.

Automaticity is different from what enables other kinds of automatic attribution. When a puppet moves when one speaks beside it the sound seems to come from the direction of the puppet. This is provided by an automatic mechanism. But in this case I can have defeating evidence, and, in fact, I do, since I know puppets do not speak. So I do not attribute speech to the puppet. I can always have defeating evidence. Evidence for the contrary will be taken into account each time for such perceptual cases which are not based solely on self-

¹⁰² One could object (like Peacocke, Wright would) to my use of the word evidence. On some views self-ascription of pain is groundless (or the subject has a default entitlement). But my idea is not dissimilar; there is some kind of asymmetry between what justifies or entitles the self-ascriptions based on knowing from the inside as opposed other-ascriptions. It cannot be groundless in a sense that whenever the subject wishes to self-ascribe property *F* then the subject is in a position to self-ascribe it. There should be correctness conditions of a self-ascription. It is not the case that self-ascription of *F* is always true or it is always true when the subject only wish to self-ascribe it. Imagining *p* and believing on this 'basis' makes the belief 'groundless'. I call 'evidence' whatever supports thinking *p* or makes self-ascription correct.

reflexive ways of knowing. In contrast, for feeling a headache through a self-reflexive way of knowing there cannot be such defeating evidence which makes me attribute the headache to someone other than me. This is why the mistake that your economist friend committed in the first case is very unusual. For normal mistakes one certainly would not consult a psychiatrist or think of irrationality. Thus the case is indeed very strange.

Through self-reflexive ways of knowing the healthy subject cannot ascribe F to anything but herself. But what enables her to know that it is of herself? Before one acquires the mastery of the first-person pronoun one knows of her body and mind already. The reason, as I understand it, is this: the content/information automatically has **non-defeasible self-significance**. Self-significance can be read off from the way the creature acts upon the information she received before she is able to use 'I'. The self-directed actions show that the creature receives self-significant information. And the non-defeasible self-significance enables the subject to use 'I' for the exposed object.¹⁰³

What does non-defeasible self-significance mean? To have non-defeasible self-significant information means that the information potentially has immediate implications for the subject's action. Self-significance should be understood in terms of a disposition to self-directed action. (The subject might not act for some reason but if the subject were to act then the information would have significance for self-directed action. This is very similar to Perry's notion of 'self-significance'.) If the subject feels thirsty, she will drink if there is a drink in the vicinity, and if the subject feels out of balance, the subject will adjust her posture. The subject might be wrong that she is thirsty or out of

¹⁰³ The self-reflexivity of ways of knowing secure that just and only the subject can be known through such special ways of knowing. But the subject might not be aware of this. Automaticity enables the subject (who is able to use 'I') to self-ascribe whatever she knows/believes through self-reflexive ways of knowing automatically. The registration of the non-defeasible self-significance (automaticity) enables the subject to self-ascribe it. This capacity of self-ascription for 'I'-users is provided by a mechanism which might break down, as Campbell (1999b) observed in the case of the syndrome called 'inserted thought'.

One might object that it is coherent even for a normal subject to think that someone else is the thinker of her own thought. Imagine a case of hypnosis or brain stimulation which makes one think something. However, such a case only affects the **autonomy** of thinking. But, in such a case, *who* is the thinker is not altered. Here, **control** over one's own thinking processes would be lost or radically diminished, but without altering who the thinker of the thought is. A less unusual case for losing control over one's own thinking processes is anger. So the notions of being in control of a thought and being the thinker of a thought should be distinguished.

balance but she cannot be wrong about who should drink or adjust her posture given that it is known through self-reflexive ways of knowing. Self-significance is the registration that the property concerns the exposed object for self-directed action guidance.

The subject need not be aware of the self-significant content in order for her self-directed action to be guided by it. The exposed object is known before one acquires the mastery of the first-person pronoun, as the one which one has to care for through self-directed actions. It seems that the self-significance of the information is registered before one acquires the use of 'I' and this enables the subject to care for itself before knowing of itself.

Passing the mirror test for self-recognition for a creature is sufficient for knowing of the exposed object as an object. Animals like elephants, apes and kids (before speech acquisition) can pass the mirror test. This is sufficient for attributing *de re* content (for elephants, apes, children) of the exposed object. This is already the registration of the self-significance of the content. The registration of self-significance is a disposition to self-directed action – a disposition toward removing the mark from the forehead.

For a subject it is possible to know of the exposed object as an object before one acquires mastery of the first-person pronoun. When a child learns the language and is only able to use her proper name for herself then it seems the child is able to think of herself as an object. When one learns to use 'I', one learns to use 'I' for the exposed object which one already knows from the inside.

Let us compare merely knowing of the exposed object with the use of 'I'. In the former case the subject only knows of the object which is herself but does not intentionally and knowingly refer to herself as herself. But knowing of herself (*de re*) already draws on the self-significance of the information received.

The mechanism that underpins automatic registration of self-significance for a way of knowing might malfunction. But in this case there will be other self-reflexive ways of knowing for which automaticity holds – and which make self-ascription possible if the subject is able to use 'I'. Proprioception, kinaesthesia, sense of balance, introspection and many other ways of knowing are self-reflexive ways of knowing. Some of them might cease to be self-reflexive ways

of knowing or self-ascriptions based on them might cease to be automatic, but still there will be always other self-reflexive ways of knowing which provide the exposed object. Knowledge of the exposed object through self-reflexive ways of knowing enables the subject to know of the object that she is without the possibility of mistake.

To sum up, automaticity is the registration of self-significance (disposition to self-directed action). The registration of self-significance enables the subject to use 'I' for the exposed object; thus it is present before one acquires mastery of the first-person pronoun.

2.6.1. Automaticity seems to be the registration of self-significance but what is the empirical evidence for it?

The registration of self-significance can be reflected in experience. One of these ways is **feeling ownership** over a limb. The positive evidence for this is the rubber hand illusion. The Rubber-hand illusion (RHI) (Botvinick and Cohen 1998; Tastevin 1937, Maravita, Spence and Driver 2003, Tsakiris 2010) appears when certain conditions are met in the standard set up. The participant places his hand on the table behind a screen, so that he is deprived of vision of this hand. The subject sees a rubber hand in an anatomically congruent position with her own unseen hand. After simultaneous stimulation on both the rubber and the real hand 75 % of the participants report a feeling of ownership over the rubber hand and they feel the touch to be where they see it. In contrast this is not the case in the non-simultaneous condition. This counts as evidence in favour of thinking that there is in some cases a positive phenomenology of a feeling of ownership over a limb. It is observed that one has no tendency to act with a limb over which one does not feel ownership. A physical threat to the rubber hand during the synchronous condition is treated by the subject as a threat to the subject herself; it induces withdrawal behaviour of the real hand (Ehrsson et al 2007). This is one way of registering self-significance, but there should be other ways of registering self-significance, like for hunger or thirst, where ownership is not involved.

Let me call the knowledge acquired through self-reflexive way(s) of knowing by automatic registration of significance, and which results in a self-

ascription: **self-reflexive knowledge**. Accordingly, what we have found is this proprioception, introspection and alike provide self-reflexive knowledge if automaticity (undefeatable registration of self-significance) obtains.

To sum up, we found four constraints which might enable the subject to know of the object which she is without the possibility of mistake. Privacy and unicity are constraints on way of knowing, while internality is a constraint on channels of information. The fourth constraint, automaticity, is a constraint on the registration of the self-significance of content gained through self-reflexive ways of knowing.

I need one more element to complete my account: the subject's referential intention. If the subject intends to refer to the exposed object known through self-reflexive ways of knowing whenever she uses 'I' then she cannot be mistaken about which object it is. But unless the subject uses 'I' for the exposed object known through a self-reflexive way of knowing when automaticity obtains, it is not clear what else could make GRR possible.

This gives us all the material that I need for articulating the Simple View. In the following section I will present the Simple View. This will be the answer to my guiding question how guarantee against incorrect reference is possible.

3. The Simple View

I am now in a position to state a broad outline of the Simple View. The Simple View is simple because it does not suppose more than that the subject uses 'I' for the object of which she knows through self-reflexive ways of knowing. The subject's use of 'I' fixes the referent of 'I' to this object. The exposed object is necessarily the subject, and, the subject, by using 'I', intends to refer to this object. Exactly this makes it possible that uses of 'I' have the guarantee against incorrect reference.¹⁰⁴

¹⁰⁴ The difference between the Simple View and Campbell's (1994a) Perceptual-Demonstrative Model for 'I' is worthy of further discussion. A useful starting point of a discussion of the differences between these two views is my distinction between first-person and demonstrative IEM. However, both views emphasise direct acquaintance for understanding the first-person pronoun. Note that the Simple View allows that introspection provides the subject with direct acquaintance with the referent of 'I' even though introspection is not a perceptual faculty. Using a Demonstrative Model for 'I' seems to commit one to denying that the cognitive significance of essential indexicals is of different kind from the cognitive significance of non-essential indexicals. I would also resist this point.

One might think that there is a simpler view than the Simple View; that is, the subject knows that 'I' stands for myself or I am referring to myself by using 'I'. I do not think this can work. It seems that one knows of the object that is the referent of 'I' from the inside and knows it in such way that one cannot be mistaken which object it is. When I am hungry I know which object my hunger concerns and I feed the right object. This cannot be explained by the knowledge that 'I' stands for myself, because this does not tell you which object you need to care about. Second the reference of 'I' and the reference of 'myself' are not fixed independently from each other. 'Myself' presupposes the capacity to understand and use 'I'. It is because either the same rule fixes the referent of 'I' and 'myself' or what fixes the referent of 'I' is involved in fixing the referent of 'myself'. Even developmentally, understanding and using 'myself' comes much later than understanding 'I' (Rochat 2003).¹⁰⁵

According to the Simple View the subject uses 'I' for the object of which the subject knows through self-reflexive ways of knowing. For a subject who uses 'I', there is always at least one way of knowing which is self-reflexive and for which the knowledge of the self-significance is automatic. Introspection or one of the bodily senses is always at hand whenever the subject is thinking and in a position to self-ascribe something.

In my view there is a single object, the exposed object which is necessarily the subject; the subject uses 'I' for the exposed object which she knows from the inside. This fixes the referent of 'I'.

4. Summary: The Simple View

Constraints	Description of the feature
1. Privacy for W (P)	Only the subject can have this way of knowing W_1 for knowing of this content, C.
2. Unicity for W (U)	Only a unique object can be known through the way of knowing.

¹⁰⁵ The capacity to use 'I' is prior and presupposed by the capacity to use 'myself'. To know what is the referent of 'I' cannot derive from another referring term the referent of which is fixed in the same way as the first-person pronoun, or which at least employs the way as the referent of 'I' is fixed.

3. Internality for channels of information (I)	The channel of information provides information, for x , of states, processes and events in x .
Self-reflexive W	The ways of knowing which are 1-3 (PIU) are self-reflexive.
4. Automaticity for registration of self-significance (A)	Automaticity is a constraint on the registration of the self-significance of content gained through self-reflexive ways of knowing.

Table 1. Constraints.

The referent of 'I' according to my proposal is fixed to that special object: the exposed object, the only one that is known through self-reflexive ways of knowing. Whenever the subject uses 'I', at least one self-reflexive ways of knowing is available to the subject. The subject simply uses 'I' for the exposed object and this fixes the referent of 'I'. This is my view. This explains how GRR is possible – why the subject cannot be mistaken about which object she is referring to by using 'I'.

Chapter 9

The Simple View: The Test

Abstract

I will argue for the Simple View by showing how it passes the Reference-fixing Test. In order to pass the test, a view must be able to account for the essential features of 'I' with the reference-fixing rule for 'I'. IEM and GRR are, of course, the key features but there are several other features as well. I will demonstrate how the Simple View can elegantly account for all of these features. Passing the test demonstrates the Simple View's explanatory power. I will argue that unless one accepts the Simple View, IEM, guaranteed right reference and aboutness-error freedom cannot be accounted for. It follows that the Simple View is the most fundamental explanation of how the referent of 'I' is fixed.

1. Preliminaries

In the literature there is a test for the reference-fixing rule of the first-person pronoun. The idea is simple. If we wish to have a theory of how the reference of the first-person pronoun is fixed, then this theory should account for all essential features of the uses of the 'I' (Peacocke 2008, Campbell 2012). I called this the Reference-fixing Test. According to Peacocke and Campbell, the key features of the uses of the first-person pronoun are IEM and GRR. There are good reasons to think they are the key features of the first person pronoun. GRR is a feature of all uses of 'I', thus the reference-fixing rule for 'I' has to account for it. IEM is explained by how the reference of the relevant kind of term is fixed (chapter 7). Thus if the Simple View is correct then it should be able to explain first-person IEM and why it is different from demonstrative IEM.

In the first part of the thesis, I already discussed the most important features of the first-person pronoun: GRR, IEM and, a further feature,

aboutness-error freedom. But a longer list of features might make our test even better. So I will expand the test to include the seven essential features listed below.

Features of 'I'	Description of the feature
1. IEM	Certain self-ascriptions made on a certain basis are immune to error through misidentification. For <i>Fa</i> based on a way of knowing <i>W</i> , who is <i>F</i> cannot be misidentified.
2. GRR	GRR is a guarantee against (i) lack of referent and (ii) incorrect reference.
3. Aboutness-error freedom	For a self-ascription, the subject cannot be mistaken about whom, which object, she is thinking about. For uses of 'I', the speaker's and the semantic referent cannot be different.
4. Error = delusion	There are certain mistakes which might count as a sign of delusion in certain conditions (e.g. "The pain which I feel is yours").
5. Semantic irreducibility (Castaneda)	The first-person pronoun cannot be reduced to other kinds of referring terms, like demonstratives, definite descriptions, proper names, or anything else. 'I' is irreducibly first personal in its semantic role.
6. In using 'I', the subject is knowingly & intentionally referring to oneself	In using 'I' 1. The subject refers to herself (<i>the exposed object</i>) 2. The subject intends to refer to herself (<i>the exposed object</i>) 3. The subject (<i>implicitly</i>) knows that she is referring to herself ($\lambda x(x \text{ refers to } x)$).
7. Functional irreducibility	Uses of 'I' have immediate consequences for the subject's action.

Table 2. Features of the first person pronoun.

I will go through each of the seven features, showing how the Simple View can account for each feature. I begin with the most important part of the test: whether the Simple View can explain first-person IEM.

2. The First Essential Feature: IEM

IEM became an important topic starting with Shoemaker because it is supposed to teach us something about the first-person pronoun. But it remains unclear what this is. Shoemaker claims that the capacity for self-ascription requires the capacity for immune self-ascription, but he does not develop the idea (Shoemaker 1968: 565-566). The Simple View will elucidate why understanding IEM is essential for understanding the first-person pronoun and provides us with an understanding of the possibility of self-ascription. So what follows is a deeper understanding of the first-person pronoun by understanding first-person IEM.

In the first part of the thesis, I argued that self-ascription – based on only knowing from the inside through self-reflexive ways of knowing – is immune relative to ‘I’ (chapter 2-8). We have found that the question of IEM of self-ascriptions, the referentiality of ‘I’, and the question of how the subject knows of the referent of ‘I’ are intimately related (chapters 1 and 2) and saw that the IEM of self-ascriptions is basis-relative (chapters 1, 3, 4, and 6). It turns out that the general explanation of IEM requires the reference-fixing rule (chapter 7). We saw that there are two different kinds of IEM: first-person and demonstrative IEM (chapter 6).

The Simple View offers connections between these seemingly unrelated segments. The referent of ‘I’ is fixed by the subject’s use of ‘I’ to the exposed object known through self-reflexive ways of knowing. When a self-ascription is based only on self-reflexive ways of knowing, then the self-ascription is immune. This is because the reference of ‘I’ is fixed through such ways of knowing. Thus without specifying the basis, we cannot evaluate whether a self-ascription is immune. For IEM it matters how the subject knows of the referent of ‘I’ because the reference is fixed through it.

I will present four questions about first-person IEM. If the Simple View is the correct view then the Simple View should help us answer these questions.

2.1. The Reference-fixing Test for the Simple View:

The explanation of first-person IEM

If the Simple View is the correct view of how the reference of 'I' is fixed, then it has to explain first-person IEM by answering the following questions:

1. What accounts for first-person IEM as a different kind of IEM?
2. Is it the case that if I encounter the same object twice, at different times, then this opens the possibility of misidentification?
3. Why are self-ascriptions immune only on a certain basis?
4. How can the Simple View account for the IEM of the self-ascriptions based on a certain way of knowing?

I will start with the first question.

2.2. What accounts for first-person IEM as a different kind of IEM?

According to the general explanation of IEM, $Fa(W)$ is immune iff the reference of 'a' is fixed through W at the time when it is fixed. However, the reference of a perceptual demonstrative expression is fixed in a different way from how the reference of 'I' is fixed. This explains the difference between first-person and demonstrative IEM.

For perceptual demonstratives, typically multiple-object ways of knowing are available to fix their reference to one of the perceived objects. Thus, for demonstrative thought the basis alone typically cannot rule out the possibility of misidentification. A self-ascription could be based on two self-reflexive ways of knowing and remain immune. In contrast, when a demonstrative thought is based on more than one way of knowing and it is fixed through one of them (e.g. vision), then it loses its IEM. I develop this in the following example.

Let us suppose that I see a dog and hear a dog barking, and based on these modalities (W_{vision} , W_{audition}) I think that dog is barking, where the referent of 'that dog' is fixed by vision. In this case I *can* misidentify which dog is barking because that which I hear might not be the one which I see. On these bases, W_{vision} , W_{audition} , misidentification is possible.

In contrast, a self-ascription that I am distressed about being out of balance is based on two self-reflexive ways of knowing (sense of balance, introspection), but it remains immune. It seems that two radically different ways of knowing can be the basis while the self-ascription remains immune. On these bases ($W_{\text{sense of balance}}$, $W_{\text{introspection}}$) misidentification is **not** possible. The object which I think of cannot be different only because it is based on two different self-reflexive ways of knowing for self-ascriptions. The Simple View accounts for this difference by acknowledging that the referent of 'I' is fixed to the exposed object which the subject knows through self-reflexive ways of knowing.

According to the general explanation of IEM, Fa (W) is immune iff the reference of 'a' is fixed through W (at the time t when it is fixed). This seems to suggest that I cannot have an immune judgment which is based on knowing of the object at different times. Is it the case that reidentification always allows for the possibility of misidentification?

2.3. Is it the case that if I encounter the same object twice, at different times, then this opens the possibility of misidentification?

I will show that first person IEM and demonstrative IEM are different because their reference-fixing rules are different. Encountering an object twice opens the possibility of misidentification for demonstrative thoughts, but not necessarily for self-ascriptions.

One might encounter an object twice and attribute F -ness based on the supposed second encounter of that object to the object which was initially encountered. However this opens the possibility of misidentification of what is F for a demonstrative thought.

Consider a case when I see a bird twice (2 minutes apart) and think I see the same bird; I have an identity presupposition (*that bird₁=that bird₂*). We have two tokens of the same type of demonstrative. Suppose that the bird looks sky blue the first time when I see a bird. In this case I can go wrong whether that bird₂ is sky blue B =vision used on two different occasions.

In contrast, when I think of myself at two different times (2 minutes apart) based on what I feel from the inside, misidentification is not possible. First I am in pain and 2 minutes later, I am no longer in pain. In thinking that I was in pain,

but it has subsided, I cannot go wrong about whose pain has subsided. (This notice of change seems to rest on an identity assumption that the one who was in pain is the one who is not in pain now.¹⁰⁶) To think that my pain has disappeared requires presupposing that I am not in pain now, and I cannot go wrong in thinking of the one who is not in pain now. And when I remember feeling pain from the inside (2 minutes ago), I cannot go wrong whose pain I am remembering. In contrast, for demonstratives, encountering an object twice in two different times opens the possibility of misidentification.

The Simple View offers a straightforward explanation for this difference. The object that the subject knows through self-reflexive ways of knowing is necessarily the subject herself. Thus when the subject uses 'I' twice at different times for the object which she knows through a self-reflexive basis, then the subject cannot go wrong. This is the case even if the basis is different but self-reflexive. In contrast, for demonstratives, such mistakes can happen if one encounters an object twice in different times – even if the judgements are made on the same basis.

The reason for this lies in the differences in how the reference of 'I' and 'that' is fixed. A perceptual demonstrative's reference is typically fixed through a multiple-object way of knowing (vision/touch/audition) to one of the many objects which is perceived.¹⁰⁷ Multiple-object ways of knowing allow that two different objects can be found in the same location at different times and the two objects might be indistinguishable. Thus, one might suppose that they are the same. It opens the possibility of misidentification when one supposes that in two different times one encounters the same object. In contrast, when the way of knowing is a self-reflexive way of knowing and one encounters an object at two different times, (putting aside quasi-memory cases) it remains the case that the object encountered cannot be anything but the subject. Consequently, what accounts for the differences between the IEM of demonstrative thought and self-ascription might well be the difference in how the reference is fixed for a demonstrative and for 'I'.

¹⁰⁶ I am setting aside quasi-memory cases. These were discussed in the Appendix to chapter 3.

¹⁰⁷ If I use a demonstrative to that object which I feel from the inside then this would not be known through multiple-object way of knowing. But such a use is atypical.

2.4. Why are self-ascriptions IEM only on a certain basis?

Self-ascriptions *are* immune iff they are based on self-reflexive ways of knowing. Through self-reflexive ways of knowing the only object which can be known is the subject.¹⁰⁸ A self-ascription not solely based on a self-reflexive way of knowing is not immune. This provides the reason why the IEM of self-ascription is basis-relative.

2.5. How can the Simple View account for the IEM of the self-ascriptions based on a certain way of knowing?

Self-ascriptions are immune iff they are based *solely* on self-reflexive ways of knowing. If a self-ascription is based on self-reflexive and non-self-reflexive ways of knowing, then it will not be immune. ('I have my legs crossed' based on proprioception and vision is not immune. The one whom I see might have her legs crossed while mine are not.) The reason why a self-ascription based only on self-reflexive ways of knowing is immune is this: the referent of 'I' is fixed to the only object that can be known through these ways. The subject uses 'I' for the exposed object which the subject knows of through self-reflexive ways of knowing. Thus whatever the subject self-ascribes based on self-reflexive ways of knowing on this basis will be immune.

What ensures that the object known through self-reflexive ways of knowing is necessarily the subject? The object known through self-reflexive ways of knowing is necessarily the object which is the subject because of the features of such ways of knowing. A self-reflexive way of knowing obeys privacy, unicity and internality. Internality guarantees that *x* receives information of *x* without supposing that *x* is a unique object (*x* might be a group or part of an object). However, unicity ensures that *x* in the second occurrence, what is known (object position) is a unique object, while privacy ensures that *x* in the first occurrence, who knows of *x* (subject position) is a unique object. Consequently, together internality, unicity and privacy ensure that what is known for the subject through self-reflexive ways of knowing is necessarily the subject.

But what could ensure that the subject necessarily intends to refer to this object and not another one? The answer is simple: the subject's referential

¹⁰⁸ There are some counterexamples to this claim which I will discuss in chapter 10.

intention is required for being able to use 'I'. It is because what fixes the referent of 'I' is this: the subject uses 'I' for the object of which she knows through self-reflexive ways of knowing. Without this 'I' cannot be used.

What would be the case if the Simple View were not true? If the referent of 'I' were not fixed to the exposed object known through self-reflexive ways of knowing, then it would **not** be necessary that the subject always uses 'I' for this object. If the subject uses 'I' for some description of herself or a demonstrated object, or a combination of these, or something else, then nothing can ensure that she necessarily uses 'I' for the object that is herself. Castaneda, Anscombe, and Perry observed that there is no description, demonstration or a hybrid of them to which the meaning of the first-person pronoun can be reduced; in some counterfactual situation they would not get the referent correct.¹⁰⁹ This indicates that a description or demonstrative **should** not be involved in how the reference of 'I' is fixed. But this raises the question of what ensures that the subject gets the object right?

The second feature in the Reference-fixing test is GRR. Here I continue our discussion with GRR.

3. Other Essential Features

3.1. GRR

Guaranteed right reference is a double guarantee for a referring expression type. Accordingly, whenever a token of this type is used it cannot lack reference (GRR1) and it cannot be used incorrectly or have mistaken reference (GRR2). It seems to be common ground in the literature that uses of 'I' have GRR.¹¹⁰

A guarantee against lack of reference requires that uses of 'I' cannot lack referent. According to the Simple View, the subject uses 'I' so unless the subject

¹⁰⁹ There are obvious candidates to answer this worry, e.g. Kaplan's Dthat. But this would only be needed if the subject cannot know of the object by direct acquaintance. There could be a definite description with rigid designation and this would get the object, the referent of 'I', right. But this might not be sufficient to enable the subject to know which object is the referent of 'I' whenever she can use 'I'. A definite description, like a descriptive name, does not provide direct knowledge of the exposed object, the referent of 'I', for the subject.

¹¹⁰ The only objector against GRR is Evans, who does not hold it for brain in the vat cases. I set such cases aside. It is unclear in such cases which creature could use 'I' if there is no subject. This is perhaps an issue for Evans. It is not clear why the ability to use 'I' wouldn't be sufficient to turn this creature into a subject.

exists 'I' cannot be used. Still the Simple View cannot claim to be privileged in accounting for the guarantee against lack of reference, because other views can account for this in a very similar manner as well. But the second guarantee is special. The Simple View can account for it by exploiting such resources which are not available for any other view.

I argued that GRR2 requires that the intended reference and the semantic reference cannot come apart for all uses of 'I'. This, as we have seen above, is just aboutness-error freedom. Thus I will answer how the guarantee against incorrect reference can be secured and explained according to the Simple View by answering how aboutness-error freedom is possible.

3.2. Aboutness-error freedom

Kripke (1977) distinguishes between speaker's referent and semantic referent. The speaker, by using a singular term, might intend to refer to a different thing than the semantic referent of that term, if there is any (it might lack reference).

Based on Kripke's idea, I worked out the notion of aboutness-error for those cases when both the speaker's and the semantic reference pick out a unique object, but they are different. When the speaker's referent (that object the speaker intends to refer to) is different from the semantic referent (that object the referring term conventionally in this context refers to) this is called an aboutness-error. Aboutness-error requires that both the speaker's and the semantic referent are individual objects but they are different. When the speaker's referent cannot be different from the semantic referent, I called this aboutness-error freedom.

As I argued, *aboutness-error* is not possible for all self-ascriptions. In using 'I', I cannot intend to use it for anyone but the object which is me.¹¹¹ The claim is that I cannot be wrong about whom I *intend* to think about or refer to because I cannot be wrong about whom I am thinking about.¹¹² I always get its referent right.

¹¹¹ "Guaranteed reference for that name "X" in this further sense [...] would entail a guarantee, not just that there is such a thing as X, but also that what I take to be X is X." (Anscombe 1975: 151).

¹¹² I argued that aboutness-error is different from misidentification. I always intend to refer to myself when I use 'I'.

It is not obvious what excludes aboutness-error. But if we accept the Simple View, aboutness-error freedom is no longer surprising. The Simple View has a simple explanation for this. According to the Simple View whenever the subject uses 'I', she uses it for the exposed object which she knows from the inside (through self-reflexive ways of knowing). In the case of a self-ascription based on self-reflexive ways of knowing, there is no way that the subject can be wrong about which object she is thinking of, because only a unique object, the exposed object which is necessarily the subject, can be known from the inside. Intending to refer to the exposed object known through such ways cannot go wrong.

Only one object can be known through these ways of knowing and the subject is automatically aware of which one it is: the primary object of care. Unless this automaticity breaks down, the subject cannot be mistaken which object it is. Whenever the subject uses 'I', at least one of the self-reflexive ways of knowing is available to the subject, enabling the subject to get the object (which she is) right. At least introspection or bodily awareness is always available to the subject, thus there is at least one way that she knows which object she is from the inside.

This is the reason why it is disturbing if I see Ceausescu's doppelganger but it is not disturbing if I see my own doppelganger. (Ceausescu – the Romanian dictator – had many doppelgangers to replace him in events which were either unimportant or too important. Ceausescu was afraid of assassination and not without reason.) I know perfectly well which object I am, the object which I know from the inside. So someone might be my doppelganger, but I have no difficulty knowing which one I am. However, I would have difficulties differentiating Ceausescu from his doppelganger.

To sum up, knowing from the inside presents the object which is the subject in such a way that the subject cannot be mistaken about which object she is. There is no other way of knowing (other than the self-reflexive ones) which can provide the subject with such certainty. So to make all self-ascriptions aboutness-error free requires that the subject use 'I' for the exposed object (known from the inside). Accordingly, the Simple View can account for the aboutness-error freedom of all self-ascriptions.

3.3. Error is delusion

The next feature, which I shall discuss, is *error is delusion*. Recall the Adorno case, when Adorno claims that he feels your pain from the inside. Call this the **error is delusion** case. This is not a mistake like others. One who is mistaken about that shows a symptom of a mental illness or cognitive disturbance. This is surprising. Usually mistakes come at a much cheaper price. A mistake usually cannot provide a reason to refer someone to a psychiatrist or be read as a sign of delusion. Something has prevented us from making such mistakes. There should be an explanation why one cannot think that what he feels is someone else's pain based on feeling. In contrast, in other cases, there is publicly available evidence and we can always try to make the evidence apparent to the other person. But this is not the case with the Adorno affair. The evidence which one has (if one has any) is not public, but private. If we speak about correctness conditions, then the correctness conditions would be tied to some internal state of the subject. The evidence should come from the inside, so if one lacks this then you cannot change his mind. What can account for this?

On the Simple View the referent of 'I' is fixed to the exposed object known through self-reflexive ways of knowing. Thus a private access to the ground or evidence behind such assertions is exactly what the Simple View predicts on self-reflexive bases. This accounts for this kind of first person/third person asymmetry that we have found in the Adorno affair.

3.4. Forms of irreducibilities

I turn to discuss other features which seem generally accepted beside GRR and IEM. The Simple View has to account for the difference between the first-person pronoun and other singular referring terms. I will discuss three different ways to spell out what the irreducibility of first-person might mean.

3.4.1. Semantic irreducibility

I turn to discuss Castaneda's point about the semantic irreducibility of the first-person pronoun.

I take this as a simple point that the meaning of the first-person pronoun cannot be reduced to other kinds of referring expressions. I label this semantic

irreducibility. Castaneda writes: “The first person pronoun used indexically is not reducible to other mechanisms of reference” (1989: 80). The replacement of the first-person pronoun, used indexically, by any co-referring expression fails to preserve truth and meaning. Let me illustrate this with some examples.

The truth of the sentence ‘I am uttering nothing’ is contingent or the sentence is at least not a flat contradiction (p and not p), as Castaneda observes. Let us replace ‘I’ with *the person uttering this token*. The truth of the sentence ‘The person uttering this token is uttering nothing’ is a flat contradiction. Thus it is necessarily false (Castaneda 1989: 208-209). I might add that rigid designation does not help. Kaplan’s ‘Dthat’ will pick up the same object in all possible worlds, thus it rigidifies the description. ‘Dthat the person uttering this token is uttering nothing’ remains a flat contradiction. With the exception of ‘I’, the subject is able to think, for all other referring expressions referring to her (e.g. a perceptual demonstrative), that this term is not referring to her (Castaneda 1988: 71-73). This is a quick illustration of the irreducibility of the meaning of the first-person pronoun.

Castaneda argued that any replacement of the token first-person pronoun used indexically with a proper name, a definite description or a demonstrative will fail for similar reasons. “The first-person pronoun, without predicating selfhood, purports to pick out a self *qua* self and (...) it invariably succeeds” (Castaneda 1966a: 90). No referring term other than the first-person pronoun can have exactly this function. Thus ‘I’ cannot be reduced to any other referring expression.¹¹³

This datum has to be accommodated. The Simple View seems to accommodate this perfectly well by pointing to the radical differences between how the reference of ‘I’ is fixed compared to other singular terms. The subject is directly acquainted with the exposed object (the referent of ‘I’) through the self-reflexive ways of knowing. Because these ways of knowing are such that only the exposed object can be known through them, the mistake of which object I know about is not possible. Only ‘I’ has aboutness-error freedom and the Simple View nicely accounts for this. I suggest that one reason why the meaning of the

¹¹³ I might add that ‘I’ is generally semantically irreducible – but not necessarily for every single token – so this allows for exceptions (see Boer and Lycan 1980).

first-person pronoun cannot be reduced to other kinds of referring expressions because none of these candidate expressions can secure aboutness-error freedom.

3.4.2. Knowingly and intentionally referring to oneself

For using 'I', the subject knowingly and intentionally refers to herself. This is supposed to capture the point which Castaneda was making according to Evans (1982) and, following him, Peacocke (2008).

When the subject uses 'I' then the subject knows that she refers to herself and the subject intends to refer to herself. This is argued to be essential for understanding the function of the first-person pronoun in our language. The only point which I shall discuss is whether the meaning or the function of 'I' can be reduced to the function of knowingly and intentionally referring to oneself.

Castaneda argued that there is a difference between two types of judgment: first-personal and non first-personal judgement. Suppose that the editor of *Soul* thinks 'The editor of *Soul* is underpaid' (J). Whatever makes the thought 'I am underpaid' (K) said by the editor true makes J true as well. Still there can be a difference. According to Castaneda, J leaves it open whether or not the speaker knows that he himself is the editor of *Soul*. For K, the editor necessarily and knowingly thinks of himself when he uses 'I', but this is not the case when he asserts J.

I can refer to myself without using 'I'. What is the difference? In using a referring expression other than 'I' for myself, I might not know that I am referring to myself. However, this is excluded for uses of 'I'. When Oedipus¹¹⁴ announces, "I will find the killer of Laius", then he does not intend to say that he will find himself. Unbeknownst to him, he is the killer of Laius and *de re* speaks of himself.¹¹⁵

¹¹⁴ Oedipus is the son of Laius (in the story of Oedipus), but at first Oedipus does not know this. He does not even know that his name is Oedipus (until the end of the story).

¹¹⁵ One possible way to capture self-ascriptions is to use *de se* content. *De se* content requires exactly the feature that is captured by knowingly and intentionally referring. Because of restrictions of space and time, I will not discuss *de se* content. I think *de se* content is a modelling device for first person content and it is a good one. But I do not think that from features of the *de se* content I have to be able to read of all the features of the first person content.

According to Evans the difference can be captured precisely by the notion of knowingly and intentionally referring to oneself (Evans 1982: 207-209). This contains three components.

The first component is this: the subject refers to herself (the exposed object).

The second component is this: the subject intends to refer to herself (the exposed object) by using 'I'.

The third component is this: the subject at least implicitly knows that she is referring to herself, i.e. $\lambda x(x \text{ refers to } x)$. Evans used the lambda operator to capture the fact that the subject knows that she is referring to herself. This makes it clear that what the subject knows is not just that under one or another mode of presentation the subject knows that she refers to herself, but the subject knows in a genuinely first personal manner that she is referring to herself $\lambda x(x \text{ refers to } x)$. The lambda operator nominalises the predicate and makes it sure that x thinks of herself under the same mode of presentation (x) as that under which she thinks that it is herself to whom he refers to (x).

The characterisation of knowingly and intentionally referring is the following:

1. The subject refers to herself (*the exposed object*)
2. The subject intends to refer to herself (*the exposed object*)
3. The subject (*implicitly*) knows that she is referring to herself ($\lambda x(x \text{ refers to } x)$).

Thus it would seem that the notion of knowingly and intentionally referring might provide the most fundamental understanding of 'I'.

My complaint is this: knowingly and intentionally referring to oneself is not sufficient for providing an explanation aboutness-error freedom and for the IEM of certain self-ascriptions. There is nothing in knowingly and intentionally referring which specifies how the subject knows of the object which she is. This view allows that one might always try to use 'I' for the wrong object – even while knowing that by using 'I' she is referring to herself, intends to refer to herself and knows that she is supposedly referring to herself.

But if one does not know which object she is one might get the object wrong. If knowingly and intentionally referring to herself would be sufficient

for using 'I' then this would allow that one might use 'I' for the wrong object. This allows for the possibility that I intend to refer to the wrong object because I think that object is me.

Why cannot I be mistaken about which object I am? How do I know of the object which I should use 'I' for? Why would I use 'I' for this object? If there were no other, more fundamental understanding of what we refer to other than the fact that we knowingly and intentionally referring to oneself in using 'I', then GRR would seem surprising. Knowingly and intentionally referring to oneself does not provide any advantage to the subject for knowing which object she is. Nothing would rule out the possibility that I think of the wrong living body as me and intend to use 'I' for this body. To rule out this possibility requires that I use 'I' for the exposed object and, thereby, for myself, as I suggest.

Recall: Wittgenstein and Anscombe were trying to understand the conditions on how the subject knows of the referent of 'I' if 'I' refers. It seemed to them that if 'I' refers then it is radically different how one knows of the referent of 'I' compared to how one knows of the referent of other kind of referring terms. I think they were correct in claiming this but wrong in thinking that it follows that 'I' does not refer. It seems that the way that the subject knows of the object which she is provides the key for understanding what differentiates the first-person pronoun from other referring terms.

I know which object I am without the possibility of mistake. When the subject uses 'I' then the subject cannot be mistaken about which object one is thinking about. When I meet my doppelganger I would have no difficulty whatsoever. If 'I' were only for knowingly and intentionally referring to myself, then I might possibly take a doppelganger of mine to be me. However, doppelgangers are safe – this cannot happen (unless one is delusional or irrational). But it would be a surprise that we cannot make such mistakes if the idea of knowingly and intentionally referring to oneself would capture the meaning of 'I' completely.

The Simple View accounts for how aboutness-error freedom is possible. Knowingly and intentionally referring to oneself would not put me in a position to know of a (physical) object which I am without the possibility of mistake

ever. But bodily self-ascriptions made on self-reflexive ways of knowing are immune. Again, this would come as a surprise if the idea of knowingly and intentionally referring to oneself would capture the meaning of 'I' completely.

This suggests that we have to modify the notion of 'knowingly and intentionally referring to oneself'. In the original formulation (Evans), the lambda operator nominalises the predicate and ensures that x thinks of herself under the same mode of presentation (x) as x knows of herself. This seems to suggest that the subject must know *explicitly* that she is self-referring. This is not necessary. On my view, this knowledge could be implicit because the self-reflexivity of ways of knowing already ensures self-reflexivity.

One might only have implicit knowledge that one is referring to oneself by using 'I' (in the third component) initially when one learns to use of 'I'. Later this knowledge becomes explicit. The subject first only *implicitly* knows that she is referring to herself ($\lambda x(x \text{ refers to } x)$); this is sufficient for using 'I'. The subject is in a position to (explicitly) know that she is referring to herself if she has the mastery of the required concepts and the ability to entertain self-reflexive thoughts. For using 'I', knowledge of self-reflexivity cannot be explicit from the beginning. It is an achievement to be able to form these thoughts. The subject simply uses 'I' for the exposed object without needing to have more than implicit knowledge of knowingly and intentionally self-referring.

In sum, it is true that when one uses 'I' then one knowingly and intentionally refers to oneself, but this is not the most fundamental understanding of the meaning of 'I'.

3.4.3. Functional irreducibility

According to Perry (1979), uses of essential indexicals ('I', 'here', 'now', 'that') have a specific functional role to play in the life of the thinker. The thinker has to be able to display the appropriate course of action related to the token indexical in the situation, including 'I'. For Perry, it is constitutive of a self-ascription, that the subject be disposed to act appropriately. An appropriate action of the subject might be some action from a range of possible actions in a suitable circumstance. Earlier Kaplan pressed a similar point. For thinking "I am on fire" the appropriate thing to do is to put out the fire for myself. To think 'I

am making a mess' is very different from *de re* thinking of myself 'that one is making a mess' without knowing that it is me. An 'I'-thought cannot be reduced to a *de re* thought of myself where I might not realise that it is of myself. This much is clear.

'I am making a mess' thought by me has limited accessibility. This thought is only accessible to me – in a first personal manner. The thought is only accessible to me when I know who is making the mess. I can understand that Caesar is making a mess even if I do not know who is Caesar. But I cannot understand that I am making a mess uttered by me unless I know who is making the mess. Everybody can believe, on some occasion, of Perry, that Perry is making a mess, but only Perry can believe 'I am making a mess' of himself when he is making a mess. This thought has limited accessibility; only Perry can access the first personal thought. Neither 'I', 'now' nor 'here' can be eliminated by other indexicals or referring expressions because of their limited accessibility.

The Simple View can accommodate Perry's point about limited accessibility for 'I'-thoughts. Only the subject can know of the exposed object through self-reflexive ways of knowing and this accounts for its limited accessibility. The way as the subject knows of that object is special and only available to the subject. This makes an 'I'-thought (like the thought that I am making a mess) only available in this way to the subject.

This is reflected, as Perry observed, in the way first-personal thoughts have immediate consequences for action as well. On the Simple View, the subject uses 'I' for the exposed object, the one under permanent care. The reason why the subject knows of the exposed object is because the subject has to take care of the exposed object in order to survive. Its self-significance marks the object under permanent care. Self-significance is a propensity for self-directed action, where a kind of self-directed outcome is the goal. For example, hunger triggers eating. The subject will eat when she is hungry and extinguish the fire when her hair is on fire because the self-significance of the content enables her to do so. On the Simple View, the exposed object known through these self-reflexive ways of knowing is already that which one has self-directed dispositions

toward. Thus, the Simple View would predict that first-personal thoughts should have immediate consequences for the subject's action.

4. IEM and GRR of self-ascriptions cannot be explained unless the Simple View is the most fundamental explanation of the first-person pronoun

We have seen that there are good reasons to accept that only the Simple View can account for all features of the first-person pronoun. For example, aboutness-error freedom, GRR and IEM are described from the subject's point of view. However if the referent of 'I' is **not** fixed from the subject's point of view, then nothing can account for this. For uses of 'I', aboutness-error freedom, IEM and GRR are features protecting the subject from certain kinds of errors. If the reference of 'I' is fixed in an objective manner (not employing the subject's perspective) then nothing can explain what provides the epistemic advantage for the subject.

If the basis of the judgment and, in particular, self-reflexive ways of knowing do not play any role in how the reference of 'I' is fixed then nothing (or it is not clear what) can explain why on some bases certain self-ascriptions are immune. It would be mere good fortune. Moreover, nothing would rule out aboutness-error. But aboutness-error is **not** possible and uses of 'I' have GRR, including a guarantee against incorrect reference. These features cannot be explained unless the referent of 'I' is fixed through the subject's point of view. We have seen how to account for all the abovementioned seven features of uses of 'I' using the Simple View. Consequently, the Simple View looks promising as an account of how the reference of 'I' is fixed.

5. Summary

In this chapter, I have discussed seven features of the use of 'I'; I have argued that the Simple View can account for all the seven features in an elegant manner. The most crucial features are IEM, GRR and Aboutness-error freedom. Neither of them can be explained unless we accept the Simple View. If the way that the reference of 'I' is fixed does not involve self-reflexive ways of knowing, then nothing can explain that only those self-ascriptions are immune which are based on self-reflexive ways of knowing. Only the subject has a guarantee

against incorrect reference and only the subject cannot misidentify who is *F* for *Fi* based only on self-reflexive ways of knowing. Thus unless the referent of 'I' is fixed – as I suggest– through self-reflexive ways of knowing then it is impossible to account for these features. Thus I think we have very strong reasons for accepting the Simple View. But to display the strength of the view I will discuss several objections in the next and final chapter.

Chapter 10

Objections to the Simple View

Objections	Source
1. How does the hearer understand the speaker's use of 'I'?	A common worry
2. The simple view is only about the psychological background, but not about the semantic question of fixing the reference of 'I'.	Fregean objection
3. <i>The Standard Objection:</i> The token-reflexive rule (TRR) for 'I' fixes the referent of 'I', so we do not need another explanation!	Defenders of TRR
4. What makes the referent of 'I' a public object when it is fixed through a private way of knowing? It seems to be a private object.	Wittgensteinian objection
5. <i>Anscombe's challenge:</i> The Simple View denies the possibility that one can use 'I' without being acquainted with herself. But Anscombe claims that a subject can use 'I' in the sensory deprivation tank without the possibility of acquaintance.	Anscombe's challenge
6. Introspection provides no object which the subject knows of.	Hume's difficulty
7. One subject might be split into different bodies, where the cognitive system might be in a body distinct from that housing the senses – this threatens my claim that only a unique object can be known through self-reflexive ways of knowing.	Dennett's and Strawson's objections
8. Asserting 'I am out of balance' based on vision and introspection is immune. But it is not based <i>only</i> on self-reflexive ways of knowing – as my explanation of the IEM of self-ascriptions would require.	Proffitt's and Shoemaker's objection
9. The notion of 'self as subject' accounts for the IEM of self-ascriptions. Nothing else is needed.	Longuenesse's objection
10. Is it possible to use 'I' without having self-reflexive ways of knowing?	Anscombe

Table 3. List of objections.

We have learned that the Simple View is, plausibly, the most fundamental explanation of how the reference of 'I' is fixed. But this is not sufficient for accepting the Simple View unless I can answer the most powerful objections against it. *The objections presented below are independent from each other, so they can be read and considered separately.*

I. General objections:

1. Objection 1:

A Worry about Understanding

The first worry is about how the hearer understands uses of 'I'; this is an obvious question that has to be settled in the beginning. Whatever way the reference of 'I' is fixed it has to be done in an intersubjective way that is available to others. Language is for communication. On the Simple View, what enables the hearer to understand the subject's uses of 'I' (where the hearer is not the subject)?

When I hear someone using 'I', I will understand that the user is using 'I' for the object which the user knows through self-reflexive ways of knowing. I take the speaker to be speaking about that object which she knows from the inside. Despite this being a simple answer, it raises some further worries.

1.1. The first worry

How does the hearer (other than the subject) understand 'I'?

The reason the hearer can understand uses of 'I' is this: the hearer knows that the subject has an intimate private access to the object which is herself. That is to say the subject knows things about herself 'from the inside' directly. This grounds the understanding of the subject. This does not require more than a basic theory of mind (Perner 1991) and, in some ways, understanding what private access is.

The basic theory of mind, in this sense, only requires knowledge that other human beings have beliefs, desires and pains (etc.) and that these are available to them in a way in which they are not available to others. The subject knows of her own thoughts, desires and beliefs in a private way. However, one knows that a good way to know of someone else's thoughts, desires, pains or intentions

is hearing them report about these things. This shows that we know in some way that the other subject has private access to his thoughts, feelings and desires. Understanding this does not require that subjects are able to articulate what they understand verbally. The subject uses 'I' in a certain way and only assumes that others use 'I' in the same way.

The subject knows how to use 'I' and assumes that others use 'I' in the same way as she does without necessarily being able to spell out the rule and how to use it. Knowing how to use the rule might consist in some form of implicit and non-propositional knowledge. I know how to walk or talk but I cannot describe how to do it in propositional terms correctly. I implicitly know the rules of language I employ but need not explicitly know the rules. The Simple View describes the practice of how the subject uses 'I' and only requires that the subject knows how to use 'I' and how to understand others' uses of 'I'.

Kripke provides a similar answer to this worry in his account of the first person pronoun, which I endorse: "If it is the sense determined by its subject's first-person acquaintance with herself, how can it be used to communicate to someone else? (...) The hearer is aware that each person, including the hearer herself, uses 'I' to refer to herself by direct self-acquaintance. Hence, knowing what it is in one's own case and taking it to be the same way for others, one understands what the first-person statement is, even though it has a sense that is, strictly speaking, incommunicable to the hearer." (Kripke 2011c: 303).

1.2. The second worry

Though Kripke answers our original question, he also presents us with a new question: Why don't I discuss the problem of sense for the first-person pronoun? 'Sense' might be understood in different ways.¹¹⁶ If 'sense' means only 'mode of presentation which solves the Frege-puzzle' (discussed briefly in the chapter 6) then, *in this sense*, my token use of 'I' has a different sense (MP) from my proper name. If 'sense' means whatever fixes the referent then I provide the sense of the first-person pronoun, but if direct acquaintance excludes that there is a sense for 'I' then I have no sense in this sense. I remain neutral on these questions. My answer depends on the notion of sense employed.

¹¹⁶ For a discussion of different notions of sense see Sainsbury (2002 and 2005).

1.3. The third worry

One may object that in a strict sense the way the referent of 'I' is fixed for a subject is not intersubjective. We agree. One could hold that whatever way the referent is fixed, it should be shared (Frege 1952),¹¹⁷ but the Simple View offers a way through which the referent is fixed which is private. In some sense this is correct. The referent is fixed through private ways of knowing. But it is shared in that for everyone it is fixed in the same way; thus everyone is in a position to understand how it is fixed for herself and for others. So this is unproblematic.

2. Objection 2:

Fregean Objection

A Fregean might worry that the Simple View is alluding to psychological processes and ways of knowing in the specification of reference and truth and claim a semantic account of the first-person pronoun should not allude to these. Thus a Fregean might claim that the Simple View only provides the psychological underpinnings or causal enabling conditions of what is constitutive at the semantic level. Accordingly, they would claim that the Simple View is a psychological explanation and this is independent of the semantic explanation.

With the exception of demonstratives and the first-person pronoun, the desire for semantic explanation without mentioning psychological processes or ways of knowing makes perfect sense. But this cannot be generalised to indexicals which require the presence of the object like 'I' or 'that' (perceptual demonstratives). Let me reformulate the Fregean objection.

2.1. The first-person perspective

The question can be put like this: could it be that the first-person point of view is necessary for fixing the reference of 'I'? (By the 'first-person point of view' I

¹¹⁷ Frege (1956) seems to agree with us and says that 'I' is an exception: the mode of presentation that fixes the referent of 'I' cannot be shared (Frege 1956: 17). There is a point of disagreement with Frege. Frege does not allow that direct acquaintance can figure into the mode of presentation (reference-fixing). I am inclined to think that the mode of presentation can be provided partially by direct acquaintance at least for the first-person pronoun (and demonstratives).

only mean the subject's point of view; it does not require the mastery of the first-person pronoun.) If the first-person pronoun has to have a shared meaning, then it is not clear how the first-person perspective can be essential for understanding it. If the first-person perspective were not constitutive of the first-person pronoun then all essential epistemic features of 'I' should appear from the third-person point of view. However this is not the case. Thus there are reasons to think that the referent of 'I' is fixed from the first-person perspective.

Nevertheless, if it turned out that the first-person perspective is not constitutive in fixing the reference of 'I' then I would be providing the psychological explanation. I doubt this. But even if this were the case, the most fundamental explanation of 'I' would be this so-called psychological one. It is because only this can accommodate the first person perspective which is essential for understanding the first person pronoun.

2.2. Answering the Fregean objection:

Third-person versus first-person point of view

Now I will try to show that the referent of 'I' *has* to be fixed from the first person perspective. This is because until I show that the reference of 'I' has to be fixed from the first person perspective, the Fregean objection does not lose its bite.

Let me give a rough idea of what the third-person perspective is. The third-person perspective is *not* from a first-person, a second-person (hearer) or a joint (we) perspective; these perspectives require the involvement of the subject. The third person perspective is, at least, a subject independent perspective like the perspective of a map. Roughly, a city on a map is presented from a third person perspective. In contrast, when I feel pain this is from the first person perspective.¹¹⁸

Call a singular term the reference of which is fixed or determined through the third-person point of view a **3P singular term**. 3P singular terms, like

¹¹⁸ Castaneda's notion of 'He*' ensures that the subject is thinking of himself as himself and introduced as a technical device to capture the first-person pronoun. 'He*' could be seen as capturing 'I' from the third-person point of view. On my view this appears to be a reductionist attempt; this might not fit the original intention of the author.

proper names, are such that they cannot have essential features requiring the first-person point of view.

Essential features of a singular term reveal the nature of a singular term. It is not merely a necessary feature, but it is constitutive for understanding *x*.¹¹⁹ The essential features of a singular term arguably derive from how their reference is fixed. For a 3P singular term the essential features would be naturally provided from the third person perspective.

If the reference of 'I' were fixed by employing the third-person perspective (as many might think) then IEM and GRR would seem miraculous. For a reference fixed from the third-person perspective no advantage for the first-person perspective can be given. So how could a use of a 3P singular term acquire such features which are features *only* from a first-person point of view? This can only be explained if the reference is *actually* fixed through the first-person point of view, even though the reference appears to be fixed through a third-person point of view. Alternatively, if we stay with the idea of 'I' being a 3P singular term, one has to argue against GRR and IEM being essential features of uses of 'I'. But possessing GRR and IEM seem to be the cornerstones of the understanding of the first-person pronoun. GRR and IEM are exactly the features which we *should* explain in order to explain the first-person pronoun. Similarly, the referent cannot be fixed from the 2nd-person or we-perspective unless the essential features of it come from the relevant perspective.¹²⁰ Consequently, 'I' cannot be a 3P singular term. Call this **radical irreducibility thesis (RIT)**.

RIT = the reference of the first-person pronoun cannot be fixed from perspectives other than the first-person perspective.

To summarise, the RIT for 'I' requires that a singular term fixed from the third person perspective cannot be endowed with features from the first person

¹¹⁹ A discussion of what it means to concern the nature of something can be found in Burge (2010: last chapter).

¹²⁰ The reference of the first-person pronoun cannot be fixed from the plural-you or we-perspective as well because its essential features are strictly restricted to the first-person point of view.

perspective. However, 'I' has such features. I have argued that the radical-irreducibility cannot be satisfied, and GRR and IEM cannot be explained, unless the referent of 'I' is fixed from the first person perspective. Thus, I have to allude to psychological processes without which self-reflexive features of special ways of knowing cannot be explained. These are required to explain how the essential features of 'I' are possible.

3. Objection 3:

The Standard Objection

The standard objection against the Simple View is that the Simple View is superfluous, because the token reflexive rule for 'I' is all what we need for fixing the referent of 'I'. This is a powerful objection which naturally pops into one's mind. There are only a few points in philosophy which everyone agrees on; one of them is that uses of 'I' refer to the speaker (Kaplan 1989). There are different versions of the token reflexive rule (TRR) for 'I'; I think they all try to capture the same idea.¹²¹ In formulating the TRR for 'I', one might use 'thinker' instead of 'speaker', because in uttering a sentence containing 'I', one might not be voicing her thought – variations are plenty. Peacocke's fundamental reference rule for 'I' (hence FRR) is one variation among many. The FRR for 'I' roughly states that a use of 'I' refers to the producer of the thinking. There are standard objections to the token reflexive rule for 'I', but they are not considered to be fatal (e.g. Bermúdez 1998).

I need not argue against the TRR for 'I' because I think that it is a good approximation for understanding 'I' and how its reference is fixed from a third-person point of view. First-person reference is fixed from a first-person point of view but it should be understandable from a third-person point of view.

¹²¹ A possible model for first-person content is *de se* thought. *De se* thought requires that the subject knows that she is referring to herself while *de re* thought about the subject only requires that the subject in fact refers to herself. Recanati criticised *de se* thought on the grounds that it cannot capture the cognitive significance reflected in IEM. Only some self-ascriptions made on certain bases are immune. Thus the cognitive significance derived from IEM is basis relative. But self-ascriptions are self-ascriptions independently of their basis. Thus *de se* thought which is a model for self-ascription independently from their bases cannot account for the cognitive significance derived from IEM. Morgan (2012) has a different argument for the same conclusion that *de se* thought and IEM are orthogonal questions.

My response to the standard objection is this: the TRR for 'I' cannot be the most fundamental explanation of how the reference of 'I' is fixed unless it can account for IEM and GRR. To show this, let us consider two plausible claims:

1. The TRR/FRR for 'I' provides the rule how the referent of 'I' is fixed (Shoemaker 1968, Peacocke 2008: 92, Campbell 2012).
2. The reference-fixing rule for 'I' should account for all essential features of uses 'I' (Peacocke 2008: 92-98, Campbell 2012).

I will call the second claim the **accountability claim**. It is a natural thought that the essential features of 'I' should be accounted for by its reference fixing rule.¹²² Consequently, our task is to examine whether the essential features of 'I' can be accounted for by employing the TRR for 'I'. I will argue that this is impossible. The accountability claim has to be met by something else – the Simple View is our natural suggestion for this purpose.

3.1. What is missing from the TRR (FRR) for 'I'?

IEM, as I argued, is explained by the way that the referent of a singular term is fixed (chapter 7). IEM is basis relative. Thus, if basis matters for IEM of self-ascriptions then it has to matter for the reference-fixing rule for 'I'. And basis matters for understanding the guarantee against incorrect reference. What could enable the subject to know which object is the speaker without the possibility of mistake (as the guarantee against incorrect reference requires)? My answer is that self-reflexive ways of knowing can ensure this. When one knows an object from the outside one always can make a mistake if one encounters this object twice. Only knowing the object through self-reflexive ways of knowing guarantee that whenever I encounter the object I will encounter the same object (setting aside quasi-memory cases).

To adopt the TRR (/FRR) for 'I' assumes that ways of knowing do not play any role in how the reference of 'I' is fixed. Without using self-reflexive ways of knowing, the guarantee against incorrect reference cannot be explained. There is no other candidate which can provide knowledge, for the subject, of the

¹²² GRR has to be accounted for because it is a feature of all uses of 'I' so whatever fixes the reference of 'I' should offer this feature as necessary for all uses of 'I'. IEM has to be accounted for because the general explanation of IEM involves the reference-fixing rule – as I discussed before.

exposed object (the subject) without the possibility of mistake. Consequently, the TRR or the FRR for 'I' is missing the basis – the way of knowing – as its constitutive part, but this is necessary for the explanation of GRR2. The accountability claim is violated.

How could the TRR for 'I' account for the IEM of basis relative judgements (chapters 4 and 6)? A self-ascription based only on proprioception or introspection is immune, but the same self-ascription based on an external way of knowing is not immune (chapters 3 and 4). The TRR for 'I' cannot explain this, because it does not rely on the basis as a constitutive part of the reference-fixing rule. IEM is basis relative but the referent of 'I' is fixed independently from ways of knowing according to the TRR and the FRR for 'I'. All basis-independent ways of fixing the referent will fail to explain how the basis of a judgment makes a difference between a judgement being immune and not being immune. Again, the accountability claim is violated. Neither IEM nor GRR can be accounted for by the TRR or the FRR.

3.2. The First-person Point of View

What is missing in the TRR/FRR for 'I' is the first person point of view. The essential features of 'I' provide epistemic advantages only for the subject (IEM, GRR2); this cannot be unless the referent of 'I' is fixed from the subject's point of view. IEM and GRR cannot be explained unless the reference of 'I' is fixed from the first-person perspective. TRR/FRR for 'I' violates both the radical-irreducibility thesis and the accountability claim. Thus, the TRR/FRR for 'I' only remains a good and useful specification of what the role of 'I' is without telling us how its reference is fixed.

Another option is that action awareness grounds the use of 'I'. What grounds the thinker/producer's knowledge of being the agent is agent awareness on this view. On views like O'Brien (2007) and Peacocke (2008), action awareness provides for the subject the knowledge of who is the thinker/producer of the subject's thought. To know that I am the thinker or the producer of the thought plausibly requires self-reflexive ways of knowing.

What matters for us is only this. If action awareness grounds the use of 'I', then it is on the condition that this kind of action awareness is only based on

self-reflexive ways of knowing. But agent awareness is not always only based on self-reflexive ways of knowing.¹²³ One might be mistaken who the agent is, but in this case the basis or channel which provides action awareness is not an internal channel or basis. Let me emphasise that the self-reflexivity of the ways of knowing matters for ensuring GRR and IEM. This is because the subject, by using 'I', cannot be mistaken about which object she is thinking of and intends to refer to.

Updating the TRR or FRR for 'I' with action awareness from the side of the subject might enable a holder of such a view to satisfy the radical irreducibility thesis but it cannot satisfy the accountability claim. Such views (adding the relevant kind of action awareness to TRR) cannot explain IEM. The basis (ways of knowing) is not part of the explanation. Such a view cannot explain how the same judgment can be immune on one basis and not immune on another basis.

3.3. TRR/FRR and semantic irreducibility

It is assumed that the meaning of the first-person pronoun is semantically irreducible. But the TRR/FRR for 'I' seems to assume that it is partially reducible, each involves a descriptive element. Thus another advantage of the Simple View is that it does not suppose that the referent of 'I' is fixed partially by a description such as 'the speaker/thinker/producer...' or that it even has any descriptive component in it. Even a direct reference theorist like Recanati assumes the TRR for 'I' as the most fundamental understanding of 'I' and explicitly states that there is a descriptive element in it (Recanati 2007: 13). This is surprising because direct reference theories (like Recanati's) were worked out in opposition to the view that the reference-fixing rule should use a descriptive component.¹²⁴ This problem can be solved, but a view which does

¹²³ We can find evidence for external cue integration in agency self-ascriptions in healthy subjects (Moore, Wegner, and Haggard 2009) and in the pantomime experiment one commits an aboutness error about who is the agent (Wegner 2002). The pantomime illusion (Wegner 2005) is as follows. The assistant of the experimenter, A, puts her hands where a different subject B's arms usually are. B does not see his own arms, just A's arms straight ahead. The experimenter gives an order "Flex your wrist!" A's wrist flexes where B's wrist usually is. B should not move his limbs at all. However, B reports that he felt that he did the wrist flexing movement to a certain extent. This is striking as B's hand was not even there where A's hands (which were flexing) were and B's hands were stationary.

¹²⁴ The only exceptions would be descriptive or stipulative names as Evans's 'Julius' discussed in chapter 7. (The referent of a descriptive name is fixed by a definite description.)

not require any descriptive element seems to be stronger. One of the major advantages of the Simple View is that it completely eliminates descriptive elements from how the referent of 'I' is fixed.

4. Objection 4:

Wittgensteinian Objection

One possible understanding of Wittgenstein's worry about IEM is as follows. If the referent of 'I' is a public (physical) object then the subject has to recognise this object when she refers to it. Recognition (of objects such as bodies) requires employment of a *non-private* way of knowing like vision.¹²⁵ But because the subject using 'I', in some cases, need not recognise the object in such a way, the referent of 'I' in this cases cannot be a public (physical) object. Wittgenstein assumes that through a private way of knowing only a private object can be known (or private access only provides a private object). Similarly, Anscombe worried that 'I' cannot refer because the subject cannot know of the material object which is the referent of 'I' for all uses of 'I'.¹²⁶ Acquaintance with a physical object through a private way of knowing is not even conceivable. The argument presupposes either that

1. nothing can be known through private ways of knowing or
2. through private ways of knowing only a private object (non-physical Cartesian Ego) can be known.

I will show that both presuppositions are mistaken.

The first option is just false, I can know of my hunger, my intention and my pain through a private way of knowing (even if this is not my only way of knowing of them). Thus something can be known through private ways of knowing.

Even Anscombe (in *Intention*) allows that the subject knows of her own physical action without observation 'from the inside' (even if one cannot use 'I' to refer). This shows that she cannot hold the first presupposition.

¹²⁵ McDowell (2013) thinks that Wittgenstein's and Anscombe's worry about whether 'I' refers is about whether we are physical objects. But this is only part of the question raised by Anscombe. The crucial question is how the subject knows of this object.

¹²⁶ Both Wittgenstein and Anscombe argued against the referentiality of 'I'. Hamilton (1991) and O'Brien (1994, 2007) have argued that even the most charitable reconstruction of the argument remains unsound. I find the subjectless views untenable and lacking any plausibility.

The second presupposition seems to me to be what is in play in the argument. Wittgenstein and Anscombe were worried that if 'I' refers then the Cartesian has an argument for the conclusion that we are Cartesian Egos. This requires the second presupposition that through a private way of knowing only a private object can be known if anything.

The problem with the second presupposition is this: even if a way of knowing is private it does not necessitate that what is known through it is a private object. When I know that my heartbeats are quick from the inside, this alone does not make the quick beating of my heart a private event. It is a public event available to the subject through a private way of knowing. The experience of it is private. But the object, my heartbeat is available to others by other means, e.g. through a stethoscope. Even for Wittgenstein one's heartbeat cannot be a private object. Thus even for him a public object can be known through private ways of knowing.

Wittgenstein might object that 'object' means object proper, not an event or a state. A proper object is fully present in any moment, it does not unfold in time and it is not a property, an event, or a state. But I am aware of my heart when I am aware of my quick heartbeat from the inside. My heart is not a private object. Consequently, knowing of something through a private way of knowing does not make the thing which I know of private. The channel or the vehicle through which I know is private but not necessarily the object which I know of.

Touch could be used as a private and as a non-private way of knowing. When I concentrate on how it feels to touch a square object then I use touch as a private way of knowing on this occasion. There are further examples of knowing a proper object through a private way of knowing. I can know of my baby or my kidney stone through touch (used as a private way of knowing) and maybe inference. Babies and kidney stones are physical objects. Thus a physical object can be known through private way of knowing.

Let me emphasise that privacy is the feature of the way of knowing and not the feature of the object.

4.1. Against Wittgenstein

Wittgenstein argued that a token 'I' in immune self-ascriptions cannot be referring because it only could refer to a Cartesian Ego – a private object. If the subject can be known through private ways of knowing then it should be a private object. This worry is unjustified for another reason. The subject's brain (or mind) already processes self-relevant information in the sense that the brain integrates information from private and non-private ways of knowing. Non-private ways of knowing are ways of knowing of multiple objects. The subject can know through non-private ways of knowing like vision, audition or touch public (physical) objects like stones, onions and people. When the brain processes information of the subject, external (vision, audition, touch) and internal information (from proprioception, kinaesthesia) is integrated (Knoblich et al. 2005, Moore, Wegner, Haggard 2009, De Vignemont forthcoming, Wong forthcoming). Thus the subject cannot be a private object because visual, audio, tactile information are integrated with information known from the inside about the subject.

Let us look at the sense of balance. There are visual cues, including the visual flow of changes in the environment relative to the movement of the subject, and this, along with balance organs in the inner ear, enables the brain to register whether the subject is in balance. So in this sense the subject need not even know that she is a physical object; her brain takes her to be a physical object. The subject's knowledge of herself is *already* based on the assumption that she is a physical object. However this leaves it open whether the subject explicitly knows that she is a physical therefore public object. This opens the possibility of error about which kind of thing we are even if the working of our brain presupposes that we are physical objects.¹²⁷

Thus from the IEM of certain self-ascriptions, one cannot argue for the claim that we are Cartesian Egos/non-physical objects.

¹²⁷ Evolution might have made a mistake. So I am not arguing for the claim that we are physical objects or essentially human animals because our brains suppose that we are physical objects. I only point to the fact that private and non-private ways of knowing are integrated.

5. Objection 5:

Anscombe's Challenge

The Simple View requires the subject's direct acquaintance with the object which is the referent of 'I' for all uses of 'I'. Anscombe poses a challenge to this with her sensory deprivation tank. In such a tank the subject has the capacity to use 'I' and think a thought like 'I was stupid to agree with this experiment'. But how can the subject be directly acquainted with the object which she is, in the sensory deprivation tank? The Simple View cannot be correct unless this is possible. Thus this is a powerful challenge to the Simple View.

The answer is straightforward. Introspection is one of several self-reflexive ways of knowing through which the subject can be acquainted with the object which is herself.

But how is knowledge of an object through introspection is possible? In order to accept introspection as an answer I have to show that an object can be known thorough introspection. This requires that I answer Hume's scepticism about introspection.

6. Objection 6:

Hume's Difficulty: Restating Anscombe's Challenge

Hume famously denied that the subject knows of herself through experience from the inside or through introspection. When he tried to find himself by directing his attention inward he did not find anything but sensations. To translate his worry against our view would amount to saying that through introspection and bodily experience (e.g. proprioception, kinaesthesia) the subject cannot be known in a direct manner. There is no object to be found. It seems to Hume that one is not acquainted with oneself though an internal way of knowing: introspection or bodily awareness. Anscombe might repeat this worry. To refer to a physical object that is present I have to be directly acquainted with this object. But one cannot be acquainted with that physical object, herself e.g. in the sensory deprivation tank. How does it come about that one can use 'I' in this situation?

I agree that introspection seems to be radically different from perception. On the face of it introspection does not seem to be directed at the exposed

object, but at states and events in the mind. However, whenever one knows a property, state or event through introspection, one is automatically in a position to self-ascribe this if one is a healthy subject. It is true that there is no perceptual encountering of the exposed object. Nevertheless, it seems clear that we can gain knowledge of ourselves through introspection. How is this possible?

Let us step back and look carefully at what we should do according to Hume. The first step is to redirect our attention inward. Where do we direct our attention? We direct it *inward*. But this seems to presuppose that the subject's aim is to know of herself by directing her attention inward. When Hume changes the direction of the attention, then he cannot direct it inward unless he knows that he is directing his attention to what he knows from the inside. The distinction between inside and outside is the shift from the subject's surroundings to the subject. Thus, even Hume knows in some way that he is knowing of himself when he turns his attention inward even if he does not feel justified to use the term 'self' for what he finds.

One might argue that the subject cannot direct his attention inward after acquiring the mastery of the first-person pronoun unless one knows that it is directed to oneself. There is no open question of the form 'Is it me?'/ 'Is it of me?' for who it is that one knows through introspection. For healthy subjects these questions are not open questions.

Let me demonstrate what I mean. Suppose I am in a Crystal building where everything is of transparent crystals. It is a building in the organic style of architecture but nearly completely transparent. I can sit on my chair if I recognise it and I can look out of the window if I find it. But I can simply look through the walls. I will see through and see the world around me. When I walk it looks like I am walking in the air. In order to see the Crystal room, crystal chair and crystal table, or the transparent crystal ice-cream, I have to direct my attention to them. I need not know it as a crystal room or crystal chair but I have to know it under some mode of presentation. If I want to see the world outside the room I only have to look through the crystal. In contrast, I cannot see the crystal room unless I direct my attention to it. I cannot direct my attention to it unless I know of it. Similarly, normally I only see the world

around me. I only pay attention to myself if I decide to do it or if something goes wrong and this captures my attention.

There is a difference between endogenous and exogenous attention. When an outside stimulus triggers the attention shift then it is exogenous. To be more precise, when the stimulus attracts attention, the attention shift is exogenous. It does not derive from the subject's decision. A sudden sharp pain might be an exogenous stimulus in a sense that it grabs attention independently from the subject's intention. But for endogenous attention the subject's intention is operative in shifting attention.

Our worry for Hume is for endogenous attention shifts. When Hume **decides** to observe his sensations and not the world around him, then he directs his attention to himself and, in some sense, unless Hume knows this, the direction shift cannot happen. For Hume, using 'I' and 'Hume' refer to the same thing. Hume has no difficulty in seeing Hume. He would not have difficulties seeing himself as himself in the mirror or seeing his hand as his hand (even while writing that he does not find himself in introspection). In this sense Hume knows of himself through external ways of knowing. Only finding himself through internal ways of knowing causes trouble. Hume does not find the object which he is in introspection. Thus, on the assumption that the referent of 'I', the self, is a physical object, in order to *try to* find the self he has to be turning his attention inward. What enables the possibility of our directing our attention inward after acquiring the mastery of 'I' is that one has to know that one turns her attention inward to herself.

But if Hume insists that inward directed attention is not self-directed, then we need an explanation of how this shift from outward to inward is possible – without the possibility of missing its target. Hume has to answer this question; the burden of proof seems to me to be on the Humeans or whoever wants to hold such a view. My explanation is simple; the ability of such an attention shift presupposes that the subject knows that she knows of herself from the inside. So it is plausible that we know that we know of **ourselves** from the inside. But how can I know *of the object* which is me from the inside?

6.1. Knowing of an object

What does it take to know of the exposed object (*de re*) or the exposed object as subject (*de se* – knowingly and intentionally referring to herself)?

The exposed object is certainly registered as object (*de re*) when it is named. Children use their proper names for themselves before using 'I' (Rochat 2003). Understanding their proper name, passing the mirror test or pointing to themselves seem to be sufficient for *de re* knowledge of the exposed object. However, *de se* knowledge requires more.

How can the subject know of the exposed object as oneself (*de se*)? What kind of acquaintance is required for this?

6.2. Hume's objection restated

One might think that what I know through self-reflexive ways of knowing is, at most, a feature or a part of me, but I am not acquainted with the object which is myself through such ways of knowing. Hume's objection might be put in this way: even if we accept the distinctiveness of self-reflexive ways of knowing this does not suffice for accepting direct acquaintance with the exposed object. How can one be acquainted with an object without discriminating or identifying that object? I will argue that this is possible.

6.3. Acquaintance

When I see an object then I am acquainted with it. Let us get clear on differences in different types of acquaintance.

There are three relevant alternative ways of seeing something which is an object.

First, I might look at an object without knowing that it is an object. Let us imagine a stonefish on a stone. (A stonefish looks just like a stone.) In this case, I might not be able to distinguish it from the stone when I see it as part of the stone. I am not acquainted with the stonefish.

Second, I might look at a stonefish as an object but think that it is a separate stone. In this case as Crane would put it I lack the canonical concept: stonefish. I know of the object (*de re*) but view the object under the wrong concept: stone.

Third, I might look at a stonefish as a stonefish. I know of the object *de re* and view it under the correct canonical concept.

Fourth, I might be looking at a giant stonefish as a stonefish in murky water. I only see parts of it as my eyes explore it but never the whole. I know of the object *de re* and view it under the correct canonical concept.

I might put the difference between these scenarios like this. In the first case, I am not acquainted with anything. In the second, I am acquainted with the stonefish as an object – it is a *de re* acquaintance with it. In the two last cases, I am acquainted with it as a stonefish. In the fourth case I am acquainted with a part and thereby with the whole. Let us call this **part-for-whole acquaintance**. There could be similar stages in being acquainted with oneself.

6.4. Answer to Hume

Being acquainted with a part is being acquainted with the whole as for a giant stonefish – part-for-whole acquaintance. We are acquainted with the Earth (construed broadly including everything which is on it).¹²⁸ What we actually experience (being awake) is always **only a part** of the Earth. But this does not prevent us from being acquainted with the Earth. We typically need not discriminate or identify the Earth. However, this is still direct acquaintance with the object. Even if I see my body most of the time I only see part of it but this is a way of being directly acquainted with it.

Let me illustrate this point. If I were living in a whale than I would have direct acquaintance with the whale around me – given that I know that I am living in a whale. Even when I see an orange I do not see its backside, thus in some sense I never see the whole object unless it is round and rotating or rolling rhapsodically (not on one axis) – but this is rare. Most of the time we see or touch only part of the object in question; it is either because our attention is restricted or because the object is too big.

When the subject knows of herself through self-reflexive ways of knowing it is similar to the kind of part-for-whole direct acquaintance which we have with the Earth and in the other examples above. When you turn your attention

¹²⁸ If Earth is not a singular referring expression then only think of the giant stonefish in murky water or the whale when one lives in it.

inward (to self-reflexive ways of knowing) you are not able to distinguish or identify an object that is yourself while you are directly acquainted with it. You are usually acquainted with it by being acquainted with a part of it or a feature of it. Through proprioception I can pay attention to my whole body (if I wish to do so) but this is unusual. The subject's knowledge of the object that she is referring to by using 'I' is provided by being acquainted with a part, a feature, or whole of it.

One might object that I could see the Earth if I were out in space or the giant stonefish by swimming away. Yes, but similarly I can see myself when I look into a mirror and I can feel my whole body from the inside through proprioception. It is possible to pay attention to my whole body and not just parts but the default would be being aware of only part of it.

Which kind of acquaintance is sufficient to turn my attention inward? *De re* acquaintance with the object which is me might be sufficient for an exogenous shift of my attention inward. But to decide (endogenous) to shift my attention inward and not to the world requires *de se* acquaintance with the object as myself. To know where to shift my attention requires paying attention to myself as myself.

Accordingly, the putative answer to Hume and Anscombe should be the following. To search in our sensations or in introspection for the whole subject is to search in the wrong place. Mostly parts or features of it can be found unless one pays attention to the whole. But to know this, the subject already has to know that she knows of the object which is the subject through such ways of knowing. Unless the subject knows this, she is not able to switch her attention intentionally from outwards to inwards in a reliable way. The subject shifts her attention inward in order to gain knowledge of the object which is herself. The subject might turn her attention to her whole body without identifying or discriminating the object. If the subject is able to gain knowledge through these ways of knowing of herself as herself then she has to know that it is of herself. If the subject did not know that she is paying attention to herself then she would not be able to gain knowledge through these ways of knowing of herself as herself. Yet it seems that a subject is *able to gain knowledge of herself* as herself

through these ways of knowing. Let me make a remark about subjectless views like Anscombe's view before answering Anscombe's challenge.

6.5. Troubles for the subjectless views of 'I'?

On some subjectless views of self-ascription all self-ascriptions are subjectless (e.g. Anscombe, Lichtenberg). For all subjectless judgements ("It is raining" or "It is snowing") there cannot be mistakes of misidentification. A dummy subject like 'it' does not afford misidentification. If self-ascriptions are all subjectless then all of them should to be immune.

This is not the case. Only some self-ascriptions are immune, but not all of them. Thus these subjectless views of 'I' cannot explain why some but not all self-ascriptions are immune. This counts as an additional reason for not accepting such subjectless views. However, subjectless views are untenable for several other commonly discussed reasons.¹²⁹

6.6. Returning to Anscombe's challenge

To answer Anscombe's challenge we have to look at what introspection requires. Introspection is only a way of knowing for the subject of her own thoughts, desires, wishes, hopes, and intentions and generally about her mental attitudes and the content of her thoughts. According to the Simple View, in the sensory deprivation tank the subject only knows of herself through introspection while other ways of knowing are shut off. It is plausible that I know of the content of my thoughts and know that they are my thoughts through introspection.

There is something special in the sensory deprivation tank. I typically can attend to the world or to myself. I am usually in a position to turn my attention to external ways of knowing (vision, audition, touch) or to internal ways of knowing (introspection, proprioception).¹³⁰ In order to employ introspection one has to direct one's attention inward and thereby to oneself. This capacity to *knowingly* turn our attention inward (endogenously) requires having the

¹²⁹ I find it plausible that the subjectless views (where 'I' is not a referring term) are untenable (Harcourt 2000, Campbell 2004, Peacocke 2008).

¹³⁰ One might turn her attention outward and to herself. Looking at my hand or my mirror image is a case of this.

implicit knowledge that one knows of herself through these ways of knowing. In the deprivation tank this attentional switch is unnecessary because no other option is available to the subject. One should not expect the subject to discriminate, to identify, to recognise or single out an object in the perceptual field when a way of knowing is not a perceptual faculty, like introspection¹³¹, or when it is a unique object perceptual faculty like proprioception.¹³²

6.7. How could I have direct acquaintance in the sensory deprivation tank?

It is only fair to ask the following question. How could introspection, for the subject, provide direct acquaintance with herself? Although the switch of attention from internal to external way of knowing is not possible in the sensory deprivation tank, introspection is still available to the subject. One might employ her memory or her introspection. This ability to switch from direct acquaintance to memory might be crucial in understanding how one can directly know of the object which she is. This switch would not be possible unless one knows how one knows of herself. What enables the subject to use introspection is the ability to know that she knows of herself through introspection. Having self-reflexive ways of knowing enable us to be directly acquainted with ourselves.

¹³¹ If introspection were a perceptual faculty then it should be understood on our model of proprioception. In this case what I hold for proprioception would be applicable to introspection.

¹³² Let me raise a worry about deprivation tanks. It is not clear that a subject in the deprivation tank is able to know ever that she is awake after long immersion. The subject might not be able to use 'I' because the shift between internal and external ways of knowing is not possible. The ability to shift between internal and external ways of knowing might be critical to staying awake. Anecdotal reports from sensory deprivation tanks in currently in use (where one floats in a fluid in the dark in silence) claim that subjects start to lose the sense of being awake; they are unable to distinguish whether they are awake or dreaming. It could feel like being in the Matrix when one does not know whether she is dreaming or not. This needs empirical support. Experiments are needed to establish this by recording brain activity during deprivation. It might turn out that we need external ways of knowing to know that we are awake and distinguish it from dreaming. In light of new empirical evidence it might turn out that the subject is not able to use 'I' in the sensory deprivation tank because she loses the capacity to be sure that she is awake – after a certain extended period of immersion. This would provide further support for my claim that the possibility of attention shift from internal to external ways of knowing is important for understanding how we know of ourselves. The capacity to employ introspection may require that the subject is in fact able to shift attention from internal to external ways of knowing.

7. Objection 7:

The Wrong Object Objection to direct acquaintance

According to the Simple View only the exposed object can be known through self-reflexive ways of knowing and the subject uses 'I' for this object. A possible objection against the Simple View is to show that the subject might only think she knows of the exposed object from the inside when this is not the case.

7.1. Strawson and Dennett

I will first describe two thought experiments and then show why the Simple View escapes the difficulty raised by them. In the first thought experiment, Strawson (in the chapter 'Persons' in *Individuals*) describes a scenario where each of the senses, as well as the brain, are placed in different bodies. There are 5 different bodies for the 5 senses and 1 additional body for the brain. Content from all senses is transferred to the brain with some transmitters. The bodies, where the senses and the brain are placed, are in 6 indistinguishable segments of our world. There is no seeming difference or mismatch between what one or the other sense gains in terms of information (even though they are experienced through different bodies). From the subject's point of view, from the inside, everything is indistinguishable from a case when all the senses are in one body. I can add a 7th body and rewire the subject's proprioception to this body. We should place it into a new indistinguishable segment of our world. The subject receives information from the 7 bodies and everything seems to him as it seems to us. The subject is ignorant about all of these rewirings. Let us forget about the feasibility and the cost of these arrangements and suppose no error in the design or in the synchronisation occurs.

If my intention cannot seemingly control what is happening I would figure out that something is illusory. So let me add that all 7 bodies move when the imaginary subject wants his body to move in exactly the same manner. Assuming no mistake happens everything will look the same as the usual set up with a unique body.

The question will be how such a subject can use 'I' for the exposed object. It seems there is more than one object which is available to her and which seems to be the exposed object. Introspection and proprioception provide different

bodies. I will argue that the rules governing the use of 'I' and the use of demonstratives for such creatures cannot be as it is for us.

In Dennett's imaginary scenario (from "Where am I?"), there is a brain in the vat (Dennett's brain) and the body controlled by it. The body is miles away from the brain. Dennett's brain in the vat controls a remote body by a sophisticated sustaining system and information channels between them. (Imagine an elaborate secret operation of the kind to save the world that is often seen on movie screens; think of this as happening in reality. When the body explodes because it is inevitable in this hazardous operation, then the brain will be hooked up with the next body.)

Remember that we are supposing that in both cases (Dennett, Strawson) the subject is ignorant about what is happening with her.

The difficulties for the Simple View, raised by the two thought experiments, are similar. It only seems to the subject that she gains knowledge from the inside (though self-reflexive ways of knowing) when this is not the case. It only appears to the subject that she knows of the exposed object through proprioception, but this is a mistake. There is no unique non-scattered object which is the exposed object. When we think of a singular referring term we think about a unique object which is the referent of such a term.

7.2. Could the patterns of inference for uses of referring expressions be intact in Strawson's and Dennett's case? Could the first-person pronoun be used?

It is plausible to think that uses of 'I' have fixed inference rules and refers to a particular physical object (Campbell 2002, 2004). A singular referring term refers to a particular object and enables us to make inferences and gather knowledge. According to the fixed reference rules, from '*Fa*' and '*Ga*' it follows '*Fa* and *Ga*'. When one thinks 'Bog (a dog) is black' and of the same dog 'Bog is barking', it follows that 'Bog is black and barking'. When there is a dog in a location and I see and hear it then I am entitled to think of it as the same dog unless there is evidence to the contrary.

Referring expressions including 'I' cannot be used in such situations if they cannot follow fixed inference patterns which are necessary for uses of them.

However, if I were Strawson's puppet if I see and hear a dog barking in the same location, then the default entitlement cannot be that I see and hear the same dog. This would be a mistake. Vision and audition are in different bodies in different segments of our world. Thus the dogs, the one viewed and the one heard, are always different.¹³³ I never ever can see and hear **the same** dog in such scenarios. Consequently, the inference rules seem invalid for Strawson's puppet. What we think about reference would be violated in these scenarios.

For example, from 'I see my nose' and 'I touch my nose', I should be able to infer that 'I touch my nose which I see'. Vision and touch are in different bodies. Thus, for Strawson's case what I see and what I touch are always different objects. When I claim I hear, touch and see myself then I will hear see and touch 3 different bodies and my cognition will be on the 4th body. It seems I would use 'I' for random objects. Uses of 'I' could not have GRR or IEM.

Thus Strawson's puppet could only use 'I*' (some term that has a superficial resemblance to 'I' but is distinct), but not 'I'. The reason for this is that the puppet's use of 'I*' does not allow for inferences, which we take for granted. This again violates basic inference rules.

In sum, on the face of it, it is not clear that Strawson's puppet can use 'I' if 'I' is a referring term referring to a particular object. It is not clear that the use of 'I*' in Strawson-like situations has the rules or epistemic constraints which are necessary for uses of 'I'. Even if in this case use of 'I*' cannot have the rules and epistemic constraints which our use of 'I' has it is not yet clear that this would support that our use of 'I' cannot have any regularity or epistemic constraint. On the contrary, it seems that Strawson's puppet cannot use 'I' only 'I*' with different rules.

Dennett's puppet will self-ascribe mental and physical predicates and attribute it to supposedly the same object. But in this situation the self-ascribed mental (psychological) and physical (material) predicates **cannot** be true of the same object. One might think that 'I*' (used in this way) is ambiguous (Stanley 1998) and have different referents and this is why the inference does not work.

¹³³ The dog which is seen is barking (a Gettier case because in the segments of world where the bodies are placed everything is the same) but the subject is not justified in thinking this.

But then 'I*' is different from 'I'; for 'I' there is no such ambiguity (Stanley 1998).

If Dennett's and Strawson's puppets were in a position to know what is happening then I think they would refrain from using 'I' and 'that' in a way as we use it. They might introduce new referring terms with new inference rules. If the puppets were informed about exactly what is happening then they would not think they know of the exposed object through proprioception. Thus this situation would not provide an objection to the Simple View.

Dennett's and Strawson's case do not question the basic inference rules for 'I' or 'that' or in general. The inference rules are sufficiently plausible or even necessary for understanding 'I' and demonstratives or even for the referentiality of referring terms (Campbell 2002). Thus it seems using 'I' is not possible for such subjects only uses of 'I*'.

However, it seems that at the end of the day I still have to consider whether it is possible that one knows of the wrong object through a purportedly self-reflexive way of knowing. I will turn to this.

7.3. Rewiring

I do not assume that whenever the subject thinks that she gets information of the exposed object through seemingly self-reflexive ways of knowing, then she, in fact, gets information of the exposed object. It might only *seem* to the subject that what she knows is of the exposed object. If the subject is, unbeknownst to herself, proprioceptively rewired with someone else, then it will only appear to the subject that she knows of herself through proprioception. But this cannot be unless the way of knowing is altered and loses its internality – and the proprioception of the other person, to whom one is rewired, loses its privacy.

Suppose that I am rewired with Mary and I have proprioceptive awareness of Mary's body and mine as well. Suppose that I judge 'I have my legs crossed' – based on proprioception only. It seems that the subject can be mistaken about which object she knows of.

The subject uses 'I' for the exposed object, but the object which seems to be the exposed object is necessarily the subject iff it is known through self-reflexive ways of knowing. But altered proprioception lacks internality.

Proprioceptive-rewiring is such a case. Thus, in such a situation, the subject can be mistaken about whether this object known through rewired-proprioception is she herself.

We should not forget that there is multisensory integration of information about one's body coming from different modalities (Knoblich et al. 2005, Moore, Wegner, Haggard 2009, De Vignemont forthcoming, Wong forthcoming). My judgment is based only on proprioception. Multisensory integration makes the possibility that I might be rewired without knowing it soon after very unlikely. If one is rewired it will be weird. If the other's legs are crossed, but not mine, I might see my uncrossed legs and feel legs, which are supposedly mine, to be crossed. It will be evident that my proprioception is no longer reliable, so, even before I realise this, my brain will stop relying on proprioception or even recognise that the channel provides information which is not about me. Thus, after the rewiring, ignorance about the rewiring is possible, but only for a very short time.

Rewired-proprioception becomes a multiple-object way of knowing and the subject will be in the position to figure this out. I initially construed the modal force of IEM as that of epistemic impossibility: misidentification is not possible, for all the subject knows from the subject's perspective.¹³⁴ (See the discussion of Strong Identification-freedom in chapter 6, part I.) Given that proprioception is rewired, it may be that misidentification is still epistemically impossible initially for the rewired subject, but it is likely that he will come to realise that proprioception is no longer a sole-object faculty in a short time. Afterwards self-ascription based on proprioception will no longer be immune.

The Simple View deals with such cases by strengthening the modal force involved in IEM, so that error through misidentification is both physically and epistemically impossible. According to the Simple View, a condition on the IEM of self-ascriptions is that the self-ascription *has to be* based on self-reflexive ways of knowing. It is not enough that the subject treats a way of knowing as

¹³⁴ Epistemic possibility can be understood on Edgington's framework, or in other frameworks – although we prefer that epistemic possibility has no logical relation to physical possibility as Edgington argues. For Edgington, a thought that *p* can be epistemologically possible for all we know, even if it is – in fact – physically not possible. We remain neutral whether logical possibility is different in kind from other notions of possibility, which Edgington denies (Edgington 2004).

self-reflexive – the way of knowing *has to actually be* self-reflexive. So, in light of this condition, if a self-ascription *Fi* is based on self-reflexive way of knowing then the subject cannot misidentify who is *F*. Consequently for judgements based on self-reflexive ways of knowing misidentification is *physically* impossible as well.

II. Objections to the Claim that the SV explains IEM

My main strategy is to show that IEM and GRR cannot be explained unless the Simple View is correct. Thus any objection to the claim that only the Simple View can explain IEM is an objection against the credibility of our reasons for accepting the Simple View. I turn to discuss such objections. First I have to raise an objection from Proffitt and Shoemaker against my own explanation of IEM.

8. Objection 8:

Proffitt-Shoemaker Objection

One may object (as the psychologist Dennis Proffitt did in response to me in conversation) to the claim that immune self-ascription could only be based on *solely* self-reflexive ways of knowing. ‘I am out of balance’ is based partially on visual clues (in Proffitt’s theory it is mainly based on visual flow – the change of visual angle and size of objects related to the subject’s movement). Similarly, Shoemaker might object that ‘I see a canary’ is immune based on vision and introspection – as this is one of his examples of IEM. This is a pressing objection against the thought that a self-ascription (B) is immune iff it is based *only* on self-reflexive way of knowing.

The answer is as follows, when I integrate visual clues and gain the content ‘I am out of balance’, the visual content which I gain through vision need not involve the subject. There are two possibilities: either I see myself or not. If the visual content includes the subject as a visible element then I can misidentify who is out of balance. In this case my judgment is not based *only* on self-reflexive ways of knowing. The one whom I see and who is out of balance might not be me. However, if the subject is not part of the visual content then the self-

ascription is immune. In this case, I do not gain information about who is out of balance through vision.

Thus for immune self-ascription, *Fi*, it is only required that the subject has to know of the object which she believes is *F* through self-reflexive ways of knowing. This might require a slight revision of the original explanation of IEM. The revised explanation of IEM of self-ascriptions is as follows:

‘I am *F*’ based on *W* is immune iff who is *F* is known only through self-reflexive *W*.

To sum up, what opens the possibility of misidentification is that the object to which I attribute *F*-ness is known through non-self-reflexive ways of knowing. But if the object to which I attribute *F*-ness is only known through self-reflexive ways of knowing then my self-ascription is immune. This can accommodate cases like Shoemaker’s example ‘I see a canary’ where the canary is known through vision and the one who sees it is known through introspection. The visual content does not include the subject.

Misidentification is possible even in this case – bases *do* matter for IEM. Imagine a case where the one who sees a canary first wins \$999999999999. I have a pair of Google glasses showing where I am and whether I see a canary. Canaries are exceedingly hard to identify so I need the Google glass to tell me that I succeeded in seeing a canary. My bases for claiming ‘I see a canary’ are introspection, vision and testimony. It may very well be that when I think I see a canary then someone else sees one and wins. After all, Google glasses are not infallible.

9. Objection 9:

Longuenesse’s Objection

If IEM of self-ascription can be explained otherwise than I explain it then my argument for the explanatory advantage of the Simple View would be radically weakened. There is an explanation of IEM that I have not discussed yet. In the Neo-Kantian tradition there is a distinction between self as subject and self as object. Beatrice Longuenesse (2012 and in her forthcoming book) uses this

distinction for understanding the IEM of self-ascriptions. Being the subject of a mental state is such that nobody except the subject can occupy that position. The subject use of 'I' derives from Kant's notion: self as subject. According to Kant all of my thoughts p_1-p_n can be amended by adding: 'I think that...' before them. Longuenesse unlike Kant only discusses self-ascriptions and employs the notion of being the subject in the explanation of IEM.

According to her, owing to the unity of consciousness, the subject of the mental state or perception is unique. This can be expressed as 'I see that p ' or 'I think that p '. Such uses of 'I' are called the subject uses of 'I', marking that it expresses who is the subject of the mental state.¹³⁵ For predicates like '...think that p ', '...feel that p ' and '...see that p ' it cannot be misidentified who is the one in the subject position. Shoemaker's example 'I see a parrot' is explained similarly. It is immune because, for the subject, a misidentification of who is the subject of her experience or thinking is not possible.

Longuenesse's idea is that whenever I think 'I think/feel/see that p ' I cannot thereby be thinking that somebody else thinks/feels/sees that p . The notion of the subject and the apprehension of 'I' cannot come into existence without presupposing the unity of consciousness. So, because of the unity of consciousness, the subject is in a position to know that she is the subject and nobody else can be in that position. For such a view the specification of the basis on which I judge that I am the subject seems not to be necessary. This is because for the subject the predicate itself (thinking/perceiving) and the unity of consciousness ensure that the subject cannot be misidentified. This is the core of how Longuenesse explains the IEM of mental self-ascriptions. This view seems to have some plausibility.

If this were the correct explanation of IEM then we do not need the Simple View and the basis of the judgment would be irrelevant for that kind of (mental) IEM. (Longuenesse allows another kind of IEM which is weaker and basis matters.) Consequently, this would undermine our strategy to argue for the SV.

However, this view has some difficulties. Consider the next example! I judge '*I see myself in the mirror*'. There are two occurrences of 'I'; the second is

¹³⁵ The self as subject in Kant is different from Wittgenstein's notion: 'I' used as a subject. For Wittgenstein we have to have a self-ascription but not for Kant.

'myself'.¹³⁶ Usually we were interested in 'the one whom I see', the second occurrence of 'I'. Let us set aside the second occurrence. We are interested in whether 'the one who sees me' can be misidentified (the first occurrence of 'I'). This ascribes *being the subject* and such according to Longuenesse has to be immune. According to Peacocke (2008) and Longuenesse only the second occurrence ('myself') can be misidentified, the first cannot be. But, in fact, the one who sees me (1st occurrence) can be misidentified. If I am looking at someone in the mirror and think that I am that one, then she has to look at me in the same way as I look at her. Consequently, if I am not seeing myself then it is someone else who I see in the mirror and **who sees me in the mirror**. If I see someone else, e.g.: Dorothea, then the one whom I see, Dorothea, looks at me through the mirror and sees me. She is the subject of her seeing me. Thus the one who sees me can be misidentified on this basis vision.

Longuenesse may respond by claiming that this is not subject use of 'I'. But then it is not clear when we have a subject use of 'I'. If this is not a subject use of 'I' then for a self-ascription like 'I see myself in the mirror' we need further information to decide when and how it counts as a subject use of 'I'. The 'I see...' seems to be insufficient to decide whether it is a subject use of 'I'. However, this was assumed in the setup of her view.

Longuenesse might reply that knowing who sees me in the mirror when it is the subject use of 'I' is based on introspection and vision. But knowing who sees me in the mirror when it is someone else is based only on vision. I am not sure. But this might be the case. If this is right then not only the subject position matters but the basis as well. We agree on this point. But according to my explanation, what matters is whether I know based on vision *who sees me*. If yes, then misidentification can happen. But then my view seems to accommodate the data better; basis matters and the basis has to be specified.

Longuenesse might offer that for each self-ascription and even for all thoughts there is a thought 'I think (as an agent) that *p*'. This might be something which we are able to know based on introspection. These attributions seem to be immune, but I would worry that they are infallible as well. There are two worries. (i) 'I think...' (where 'think' means being the agent

¹³⁶ I follow Peacocke in assuming that 'myself' is an occurrence of 'I' (Peacocke 2008).

of the thinking) can be attached to any thought, not only to self-ascriptions, (ii) following this procedure the thought 'I think that p' will be infallible. Infallibility will become the explanatory target. From infallibility IEM follows. (This violates the self-sufficiency principle discussed in chapter 5; IEM becomes a derivative feature.)

Longuenesse distinguishes mental and bodily self-ascriptions and argues that mental self-ascriptions have a stronger IEM. Yet there is a gap between whether infallibility or immunity matters. Why would IEM be the explanatory target when it only follows from infallibility?

10. Objection 10:

Is it possible to use 'I' without having self-reflexive ways of knowing?

I have to settle one last objection which some might raise at this point. There could be a creature without self-reflexive ways of knowing, but this creature seems to be able to use 'I'. This creature is able to take care of itself; the creature *de re* knows of itself and is not under systematic delusion/error. This is an objection because it is incompatible with the Simple View.

I will argue that this creature only seems to be able to use 'I', but effectively it cannot use 'I'. They can only use a referring term which is a proper name because it lacks GRR. Let us consider *a simplified* version of Anscombe's thought experiment about 'A'-users (Anscombe 1975). 'A' is a referring term similar to 'I' but its reference is fixed differently. 'A' purports to refer to the subject but in the following way. The 'A'-user uses 'A' *only* for the foreshortened object that the user sees when she looks down. 'A'-users cannot use 'I' and only know of their body by perceptual observation or testimony. I add that they cannot use 'A' for the object they know through proprioception and other self-reflexive ways of knowing. They are deprived of all self-reflexive ways of knowing.¹³⁷

'A'-users do not know of the exposed object from the inside. One could be mistaken about whom one sees. Thus, 'A'-users can be mistaken about which object they use 'A' to refer to. If an 'A'-user looks down she might see another body than her own body. Such a subject knows through multiple-object ways of

¹³⁷ According to Anscombe proprioception is non-observational but 'A'-users would become 'I'-users if they were using 'I' for the exposed object. Thus we have to rule this out.

knowing of the referent of 'A'. Thus nothing will rule out the possibility that 'A' is used for the wrong object.

It is generally assumed that uses of 'I' have GRR and uses of 'I' are different from uses of proper names (or demonstratives) which can purport to refer but fail. Uses of 'A' would lack GRR and would never be immune on any basis. 'A' would work very similarly to a proper name in the respect that it can purport to refer, but fail to do so. A subject which lacks self-reflexive ways of knowing cannot refer to herself without facing the possibility of aboutness-error. An 'A'-user can be mistaken about which object she is. Uses of 'I' are aboutness-error free. Thus such a subject cannot use 'I'.

Conclusion

1. The Simple View

I have argued that the Simple View provides the most fundamental understanding of how the reference of 'I' is fixed. I will rehearse the basics of the Simple View. According to the Simple View the subject knows of the object, which she is, through self-reflexive ways of knowing. Self-reflexive ways of knowing satisfy privacy, unicity and internality (chapter 8).

The first constraint, privacy, is a constraint on a kind of way of knowing *W*. It ensures that only the subject can gain knowledge of the particular state of affair in question through this kind of way of knowing *W*. For example, proprioception only provides content about the subject. Thus proprioception is a private way of knowing relative to all contents which can be gained through this kind of *W*.

The second constraint, unicity, is a constraint on a kind of *W* ensuring that only a single object can be known through this kind of *W*. The third constraint, internality, is a constraint on a channel of information. It ensures that *X* receives information of *X* – while *X* might refer to a group, an individual, or parts of the organism. Some channels of information of the subject are ways of knowing. Proprioception, kinaesthesia, nociception and introspection are the most obvious examples of self-reflexive channels of information which are ways of knowing. The object which is known through self-reflexive ways of knowing is necessarily the subject. Self-reflexive ways of knowing satisfy privacy, unicity and internality. Privacy ensures that such a way of knowing is only available for the subject, *S*. Unicity requires that only a unique object *Q* can be known through it; and internality ensures that $S=Q$. Thus, necessarily the subject, *S*, knows of *S* through self-reflexive ways of knowing. Consequently, through self-reflexive ways of knowing only the subject can be known.

The subject is not yet in a position to know that she knows of herself through self-reflexive ways of knowing. But such contents which are acquired through self-reflexive ways of knowing have self-significance before the subject

acquires the mastery of the first-person pronoun. When content has self-significance then the subject has a disposition to act in a self-directed manner, e.g. when it feeds itself (X acts on X).

When I say 'I combed my hair' the action is self-directed. Eating, drinking and scratching have a self-directed structure (X act on X). Feeling hungry, thirsty, or itchy dispose the subject to act on the subject – it is a selection for self-directed action. The first-person pronoun is used for the exposed object which one knows from the inside through self-reflexive ways of knowing. The subject may use her own proper name for the exposed object, but the subject might make a mistake whose name it is. In contrast, mistake concerning which object the subject uses 'I' for is not possible. The exposed object (known through self-reflexive ways of knowing) is necessarily the subject. 'I' is used for this object.

2. The Referentiality-test

It is easy to propose a view. What is hard is deciding whether it is correct. My strategy was to use the Reference-fixing test from Peacocke (2008 and 2012) and Campbell (2012) to test the Simple View. The idea is straightforward. If we wish to have a theory of how the reference of the first-person pronoun is fixed, then this should account for all essential features of the uses of the 'I'. If uses of 'I' have essential features, like GRR, and some self-ascriptions being IEM, then the reference-fixing rule for 'I' cannot be such that it renders an essential feature to be contingent. Thus the reference-fixing rule cannot allow a failure of GRR.

The explanation of IEM teaches us that IEM is a consequence of how a singular term's reference is fixed and how the subject gains knowledge of the object which is the referent (chapters 1, 4, 6, 7). The reason we have immune thought, *Fa* (W), is, I suggest, that *a*'s reference is fixed through W. I provided several reasons for accepting that first-person IEM and demonstrative IEM are different kinds of IEM (chapter 6). However what accounts for their differences are the differences in their reference fixing-rule (chapter 9).

3. Passing the Referentiality-test

My positive argument for the Simple View is to show that it passes the Referentiality-test. The Simple View enables us to account for how a guarantee against incorrect reference is possible (chapter 9). The guarantee against incorrect reference requires that the subject cannot intend to use 'I' for the wrong object; the intended and the semantic referent of 'I' cannot be different (chapter 2). At first sight this seems to be impossible – or even mystical. For all singular referring terms it seems I could be ignorant about what its referent is except for my use of the first-person pronoun. If the subject did not use 'I' for the exposed object then nothing could explain what excludes such a mistake.

The Simple View also enables us to account for how IEM of self-ascriptions made on a certain basis is possible. I argued that IEM of self-ascriptions is basis-relative (chapters 3-4) which allows that the basis (way of knowing) plays a role in the reference-fixing of 'I'. Made on a certain basis, a self-ascription is immune because this is one of the self-reflexive ways of knowing and the referent of 'I' is fixed through such ways of knowing (chapter 9).

The basis, way of knowing, is not mentioned in the TRR for 'I', thus it is not possible that the TRR for 'I' can explain why IEM of self-ascriptions is basis-relative. However the Simple View states that 'I' is used for the object which is known through self-reflexive ways of knowing. Thus the Simple View can account for and would predict the basis-relativity of the IEM of self-ascriptions.

In view of the range of aspects of the first person pronoun it can explain, the Simple View is, plausibly, the best account we have of how the reference of 'I' is fixed.

Bibliography

- Almog, J., Perry, J., and Wettstein, H. (eds.) 1989, *Themes from Kaplan* (Oxford: Oxford University Press)
- Alston, W.P. ,1971, 'Varieties of privileged access', *American Philosophical Quarterly* 8: 223–241.
- Anscombe, G.E.M., 1963/1994, 'The First Person' in Cassam (ed.) *Self-Knowledge* (Oxford: Oxford University Press)
- Anscombe, G.E.M., 1972, *Intention* (Oxford: Basil Blackwell)
- Bach, K., 1987, *Thought and Reference* (Oxford: Clarendon Press)
- Bermúdez, J.L., 1998, *The Paradox of Self-Consciousness* (Cambridge, MA: MIT Press)
- Bermúdez, J.L. (ed.) 2005, *Thought, Reference and Experience: Themes from the Philosophy of Gareth Evans* (Oxford: Clarendon Press)
- Bermúdez, J.L. Marcel, A., and Eilan, N. (eds.) 1995, *The Body and the Self* (Cambridge: MIT Press)
- Boer, S., and Lycan, W., 1980, 'Who, Me?' *The Philosophical Review* 89: 427-466
- Bottini, G., Bisiach, E., Sterzi, R., and Vallar, G. 2001, 'Feeling touch in someone else's hand' *NeuroReport* 13: 249-252
- Burge, T., 1973, 'Reference and Proper Names' *The Journal of Philosophy* 70: 425-439
- Burge, T., 1977, 'Belief De Re' *The Journal of Philosophy* 74: 338-362
- Burge, T., 2010, *Origins of Objectivity* (Oxford: Clarendon Press)
- Campbell, J., 1994a, *Past, Space and Self* (Cambridge Mass.: Bradford Books, MIT Press)
- Campbell, J., 1994b, 'Self-consciousness and 'I'' in Campbell *Past, Space and Self* (Cambridge Mass.: Bradford Books, MIT Press)
- Campbell, J., 1999a, 'Immunity to error through misidentification and the meaning of a referring term' *Philosophical Topics* 25: 89-104
- Campbell, J., 1999b, 'Schizophrenia, the Space of Reasons and Thinking as a Motor Process' *Monist* 82: 609-625
- Campbell, J., 2002, *Reference and Consciousness* (Oxford: Clarendon Press)
- Campbell, J., 2004, 'What Is To Know What 'I' Refers to?' *The Monist* 87: 206-218

- Campbell, J., 2007, 'Demonstrative Reference, the Relational View of Experience and the Proximity Principle' in Jeshion (ed.) *New Essays on Singular Thought* (Oxford: Oxford University Press)
- Campbell, J., 2009, 'Consciousness and Reference' in McLaughlin, Beckermann and Walter (eds.) *Oxford Handbook of Philosophy of Mind* (Oxford, Oxford University Press.): 648-662.
- Campbell, J., 2012, 'On the Thesis that 'I' is not a Referring Term' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Carston R., 2002, *Thoughts and Utterances* (Oxford: Blackwell)
- Cassam, Q. (ed.) 1994, *Self-Knowledge* (Oxford: Oxford University Press)
- Cassam, Q. 1997, *Self and the World* (Oxford: Calderon Press)
- Cassam, Q., 2004, 'Introspection, Perception and Epistemic Privilege' *The Monist* 87: 255-274
- Cassam, Q., 2007a, 'Ways of knowing', *Proceedings of the Aristotelian Society*
- Cassam, Q., 2007b, *The Possibility of Knowledge* (Oxford: Clarendon Press)
- Castaneda, H.-N., 1966a, 'On the Phenomeno-Logic of the I' in Castaneda *The Phenomeno-Logic of the I* (Bloomington: Indiana University Press)
- Castaneda, H.-N., 1966b, ' 'He': a Study in the Logic of Self-Consciousness' *Ratio* 8: 130-157, reprinted in Castaneda *The Phenomeno-Logic of the I* (Bloomington: Indiana University Press)
- Castaneda, H.-N., 1966c, *The Phenomeno-Logic of the I* (Bloomington: Indiana University Press)
- Castaneda, H.-N., 1989, i (Minneapolis: University of Minnesota Press)
- Chisolm, R., 1981, *The First Person* (Sussex: Harvester Press)
- Coliva, A., 2006, 'Error Through Misidentification: Some Varieties' *The Journal of Philosophy*
- Coliva, A., 2012, 'Which 'key to all mythologies' about the self? A note on where the illusions of transcendence come from and how to resist them' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)

- Crane, T., 2003, 'The intentional structure of consciousness' in Smith and Jokic (eds.) *Consciousness: New Philosophical Perspectives* (Oxford: Oxford University Press)
- Crane, T., 2011, 'The Singularity of Singular Thought' *Proceedings of the Aristotelian Society Supplementary Volume LXXXV*
- Crane T. (ed.) 1992, *The Contents of Experience* (Cambridge: Cambridge University Press)
- Dennett, M., 1978, *Brainstorms* (Montgomery, Vt: Bradford Books)
- Dennett, M., 1978, 'Where am I?' in Dennett *Brainstorms*
- Diamond, C., and Teichman, J. (eds.) 1979, *Intention and Internationality: Essays in Honour of G. E. M. Anscombe* (New York: Cornell University Press)
- Dickie, I. 2007, 'We are Acquainted with Ordinary Things' in Jeshion (ed.) *New Essays on Singular Thought* (Oxford: Oxford University Press)
- Donnellan, K., 1966, 'Reference and Definite Description' *The Philosophical Review* 77: 281-304
- Donnellan, K., 1970, 'Proper Names and Identifying Descriptions' *Synthese* 21: 335-358
- Dretske, F., 1969, *Seeing and Knowing* (London: Routledge & Kegan Paul)
- Dretske, F., 1997, *Naturalising the Mind* (Cambridge Mass.: Bradford Books, MIT Press)
- Edgington, D 2004, 'Two kinds of Possibility' *Proceedings of the Aristotelian Society, Supplementary Volume* 78: 1-22
- Ehrsson, H. H., Wiech, K., Weiskopf, N., Dolan, R. J. 2007, 'Threatening a rubber hand that you feel is yours elicits a cortical anxiety response' *Proceedings of the National Academy of Sciences* 104: 9828-9833
- Evans, G., 1973, 'The Causal Theory of Names,' *Proceedings of the Aristotelian Society, Supplementary Volume* 47: 187-208
- Evans, G., 1982, *The Varieties of Reference* (Oxford: Oxford University Press)
- Frege, G., 1952, 'On Sense and reference' in Geach and Black (eds.) *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell), pp. 57-78

- Frege, G., 1956, 'The Thought: A Logical Inquiry', reprinted in Ludlow *Readings in the Philosophy of Language* (Cambridge Mass.: Bradford Books, MIT Press)
- Frith, C.D., 1992, *The Cognitive Neuropsychology of Schizophrenia* (Hillsdale: Lawrence Erlbaum Associate, Publishers)
- Frith, C.D., 2005, 'The self in action: Lessons from delusion of control' *Consciousness and Cognition* 14: 752–770
- Frith, C.D. Blakemore, S. and Wolpert, D., 2000, 'Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action", *Brain Research Reviews* 31
- Fulford, B., Davies, M., and Gipps, R. (eds.) 2014, *The Oxford Handbook of Philosophy and Psychiatry* (Oxford: Oxford University Press)
- Gallois, A., 2000, 'First Person Accessibility and Consciousness' *Philosophical Topics* 28
- Geach, P., and Black, M. (eds.) 1977, *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell)
- Graham, G., and Stephens, L., 2000, *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts* (MIT Press)
- Grahek, N., 2001, *Feeling Pain and Being in Pain* (Cambridge Mass.: Bradford Books, MIT Press)
- Hamilton, A., 1991, 'Anscombe and Cartesian Scepticism' *Philosophical Quarterly* 41: 39-54
- Harcourt, E., 2000, 'The first person: problems of sense and reference' *Royal Institute of Philosophy Supplement* 46: 25-46
- Higginbotham, J., and Folescu, M. 2012, 'Two takes on the de se' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Hornsby, J. 1977, 'Singular Terms in Contexts of Propositional Attitude', *Mind* 86: 31-48
- Hume, D., 1888/1978, *The Treatise of Human Nature*, Selby, Bigge, Nidditch (eds.) (Oxford: Clarendon Press).
- Hyman, J., and Steward, H. (eds.) 2004, *Agency and Action* (Cambridge: Cambridge University Press)

- Ismael, J., 2012, 'Immunity to error as an artefact of transition between representational media' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Jeshion, R., 2007, 'Singular Thought: Acquaintance, Semantic Instrumentalism and Cognitivism' in Jeshion (ed.) *New Essays on Singular Thought* (Oxford: Oxford University Press)
- Jeshion, R. (ed.) 2010, *New Essays on Singular Thought* (Oxford: Oxford University Press)
- Kanwisher, N., McDermott, J., and Chun, M.M., 1997, 'The fusiform face area: a module in human extrastriate cortex specialized for face perception' *Journal of Neuroscience* 17: 4302-43011
- Kaplan, D., 1989, 'Demonstratives' in Almog and Perry (eds.) *Themes from Kaplan* (Oxford: Oxford University Press)
- Knoblich, G., Thornton, I., Grosjean, M., and Shiffrar, M., 2005, *Human Body Perception from the Inside Out* (New York and Oxford: Oxford University Press)
- Kripke, S., 1977, 'Speaker's Reference and Semantic Reference', French et al. (eds.) *Midwest Studies in Philosophy* 2: 255-276
- Kripke, S., 1979, 'A Puzzle about Belief', in Margalit (ed.) *Meaning and Use* (Dordrecht: Reidel)
- Kripke, S., 1980, *Naming and Necessity* (Oxford: Blackwell)
- Kripke, S., 2011a, *Philosophical Troubles* (Oxford: Oxford University Press)
- Kripke, S., 2011b, 'Frege's theory of Sense and Reference: Some Exegetical Notes' in Kripke *Philosophical Troubles* (Oxford: Oxford University Press)
- Kripke, S., 2011c, 'The First Person' in Kripke *Philosophical Troubles* (Oxford: Oxford University Press)
- Lane, T., and Liang, C., 2011, 'Self-Consciousness and Immunity' *The Journal of Philosophy*, February 78-99
- Lewis, D., 1979, 'Attitudes De Dicto and De Se' *Philosophical Review* 88 reprinted in Lewis *Philosophical Papers*, Volume I (Oxford: Oxford University Press)
- Locke, J., 1694/1975, *An Essay Concerning Human Understanding*, Nidditch (ed.) (Oxford: Clarendon Press)

- Longo, M. R., and Haggard, P., 2012, 'What Is It to Have a Body' *Current Directions in Psychological Science* 21: 140–145
- Longuenesse, B., 2012, 'Two uses of 'I' as subject?' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Longuenesse, B., Forthcoming 'Kant and Freud on 'I'' *Proceedings of the 11th Kant Congress*
- Ludlow, P. (ed.) 1997, *Readings in the Philosophy of Language* (Cambridge Mass.: Bradford Books, MIT Press)
- MacDonald, C., Smith, B., and Wright, C. (eds.) 1998, *Knowing Our Own Minds* (Oxford: Clarendon Press)
- Margalit, A. (ed.) 1979, *Meaning and Use* (Dordrecht; Reidel)
- Martin, M.G.F., 1992, 'Sight and Touch', in Crane (ed.) *The Contents of Experience* (Cambridge: Cambridge University Press)
- Martin, M.G.F., 1995, 'Bodily Awareness: a Sense of Ownership', in Bermúdez et al. (eds.) *The Body and The Self* (Cambridge, MA: MIT Press)
- Martin, M.G.F., 2002, 'Singular Thought and Particular Thoughts' in O'Hear (ed.) *Thought and Language* (Cambridge: Cambridge University Press)
- Matthen, M. (ed.) 2014, *The Oxford Handbook of Perception* (Oxford: Oxford University Press)
- McCulloch, G., 1989, *The Game of the Name: Introducing Logic, Language and Mind* (Oxford: Clarendon Press)
- McDowell, J., 1984, 'De Re senses' in Wright (ed.) *Frege: Tradition & Influence* (Oxford: Blackwell)
- McDowell, J., 1994a, *Mind and World* (Cambridge: Harvard University Press)
- McDowell, J., 1994b, 'The Content of Perceptual Experience' *Philosophical Quarterly* 44: 190-205
- McDowell, J., 1986, 'Singular Thought and the Extent of Inner Space' in Petit and McDowell (eds.) *Subject, Thought and Content* (Oxford: Clarendon Press)
- McDowell, J., 2013a, *The Engaged Intellect* (Cambridge: Harvard University Press)
- McDowell, J., 2013b, 'Referring to Oneself' in McDowell *The Engaged Intellect* (Cambridge: Harvard University Press)

- McLaughlin, B, Beckermann, A., and Walter, S. (eds.) 2009, *Oxford Handbook of Philosophy of Mind* (Oxford, Oxford University Press,)
- Montague, R. 1960, 'Logical Necessity, Physical Necessity, Ethics and Quantifiers' *Inquiry* 3: 259-269
- Moore, A., 1997, *Points of View* (Oxford: Oxford University Press)
- Moore, J., Wegner, D., and Haggard, P., 2009, 'Modulation the sense of agency with external cues' *Consciousness and Cognition* 18: 1056-1064
- Morgan, D., 2012, 'Immunity to error through misidentification: what does it tell us about de se?' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Nunberg, G., 1993, 'Indexicality and Deixis' *Linguistics and Philosophy* 16: 1-43
- O'Brien, L., 1994, 'Anscombe and the Self-Reference Rule' *Analysis* 54: 277-328
- O'Brien, L., 1995a, 'The Problem of Self-Identification' *Proceedings of the Aristotelian Society* 95: 235-251
- O'Brien, L., 1995b, 'Evans on Self-Identification' *Nous* 29: 232-247
- O'Brien, L., 1996, 'Solipsism and Self-Reference' *European Journal of Philosophy* 4:176-196
- O'Brien, L. 2007, *Self-Knowing Agents* (Oxford, New York: Oxford University Press)
- O'Brien, L., 2012, 'Action and immunity to error through misidentification' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- O'Hear, A. (ed.) 2002, *Thought and Language* (Cambridge: Cambridge University Press)
- Olson, E. T., 1997, *The Human Animal, Personal Identity Without Psychology* (Oxford: Oxford University Press)
- Parfit, D., 1984, *Reasons and Persons* (Oxford: Oxford University Press)
- Peacocke, C.A.B., 1983, *Sense and Content* (Oxford, Calderon Press)
- Peacocke, C.A.B., 1992, *A Study of Concepts* (Cambridge, MA: MIT Press)
- Peacocke, C.A.B., 1999, *Being Known* (Oxford: Oxford University Press)
- Peacocke, C.A.B., 2008, *Truly Understood* (Oxford: Oxford University Press)
- Peacocke, C.A.B., 2008/2009, Self-Knowledge and Self-Reference lectures in UCL

- Peacocke, C.A.B., 2012, 'Explaining de se phenomena' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Peacocke, C.A.B., 2013 'Perception and the First Person' in Matthen (ed.) *The Oxford Handbook of Perception* (Oxford: Oxford University Press)
- Peacocke, C.A.B., 2014, *The Mirror of the World* (Oxford: Oxford University Press)
- Perner, J., 1991, *Understanding the Representational Mind* (Cambridge, MA: MIT Press)
- Perry, J., 1979/1993, 'The Problem of the Essential Indexical' in Perry *The Problem of the Essential Indexical and Other Essays* (Oxford: Oxford University Press)
- Perry, J., 1986, *The Problem of the Essential Indexicals and Other Essays* (New York: Oxford University Press)
- Petit, P., and McDowell, J. (eds.) 1986, *Subject, Thought and Content* (Oxford: Clarendon Press)
- Prosser, S., 2012, 'Sources of immunity to error through misidentification' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Pryor, J., 1999, 'Immunity to error through misidentification' *Philosophical Topics* 26/1-2
- Pylyshyn, Z., 2007, *Things and Places: How the Mind Connects to the World* (Cambridge, MA: MIT Press/Bradford Books)
- Recanati, F., 1993, *Direct Reference: From Language to Thought* (Oxford: Basil Blackwell)
- Recanati, F., 2007a, 'In Defence of Acquaintance' in Jeshion (ed.) *New Essays on Singular Thought* (Oxford: Oxford University Press)
- Recanati, F., 2007b, *Perspectival Thought* (Oxford: Oxford University Press)
- Recanati, F., 2012, 'Immunity to error through misidentification: What it is and where it comes from' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Recanati, F., 2013, *Mental Files* (Oxford: Oxford University Press)

- Recanati, F., and Prosser, S. (eds.) 2012, *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- Rochat, P., 2003, 'Five levels of self-awareness as they unfold early in life' *Consciousness and Cognition* 12: 717-731
- Roessler, J., 1999, 'Perception, Introspection and Attention', *European Journal of Philosophy* 7: 47-64
- Roessler, J., 2003, 'Intentional Action and Self-Awareness' in Roessler and Eilan (eds.) *Agency and Self-Awareness* (Oxford: Oxford University Press)
- Roessler, J., 2014, 'Thought insertion, self-awareness, and rationality' in B. Fulford, M. Davies & R. Gipps (eds.) *The Oxford Handbook of Philosophy and Psychiatry* (Oxford: Oxford University Press)
- Roessler, J., and Eilan, N. (eds.) 2003, *Agency and Self-Awareness* (Oxford: Oxford University Press)
- Rorty, A. (ed.) 1976, *The Identities of Persons* (Berkeley: University of California Press)
- Rumfitt, I., 2005, 'Plural Terms: Another Variety of Reference?' in Bermúdez (ed.) *Thought, Reference and Experience: Themes from the Philosophy of Gareth Evans* (Oxford: Clarendon Press), pp. 84-123
- Russell, B., 1905, 'On Denoting' *Mind* 14: 479-493
- Russell, B., 1912, *The Problems of Philosophy* (Oxford: Oxford University Press).
- Russell, B., 1918, 'Knowledge by Acquaintance and Knowledge by Description, in Russell *Mysticism and Logic and Other Essays* (London Longmans, Green and Co.), pp. 209-232
- Russell, B., 1918, *Mysticism and Logic and Other Essays* (London Longmans, Green and Co.)
- Sacks, O., 2010, *The Mind's Eye* (London: Picador)
- Sainsbury, R. M., 2002, *Departing from Frege: Essays in the philosophy of language* (London, New Your: Routledge)
- Sainsbury, R. M., 2005, *Reference without Referents* (Oxford: Clarendon Press)
- Salmon, N., 1986, *Frege's Puzzle* (Cambridge MA: MIT Press)
- Scholl, B (ed.) 2002, *Objects and Attention* (Cambridge: Cambridge University Press)

- Shoemaker, S., 1963, *Self-Knowledge and Self-Identity* (Ithaca: Cornell University Press)
- Shoemaker, S., 1968, 'Self-Reference and Self-Awareness', *The Journal of Philosophy* 65:555-567
- Shoemaker, S., 1984a, *Identity, Cause and Mind: Philosophical Essays* (Cambridge: Cambridge University Press)
- Shoemaker, S., 1984b, 'Persons and their Pasts' in Shoemaker *Identity, Cause and Mind: Philosophical Essays* (Cambridge: Cambridge University Press)
- Shoemaker, S. 1984c, 'Personal identity: A materialist account' in Shoemaker and Swinburne *Personal Identity* (Oxford: Blackwell)
- Shoemaker, S. 1986, 'Introspection and the Self' in French et al. (eds.) *Studies in Philosophy of Mind, Midwest Studies in Philosophy* 10: 101-120
- Shoemaker, S., 1996a, *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press)
- Shoemaker, S., 1996b, 'Self-knowledge and "Inner Sense"' in Shoemaker *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press)
- Shoemaker, S., and Swinburne, R. 1984, *Personal Identity* (Oxford: Blackwell)
- Smith, Q., Jokic, A. (eds.) 2003, *Consciousness: New Philosophical Perspectives* (Oxford: Oxford University Press)
- Snowdon, P., 2014, *Persons, Animals, Ourselves* (Oxford: Oxford University Press)
- Stanley, J., 1998, 'Persons and their Properties', *Philosophical Quarterly* 48
- Strawson, P. F., 1959, *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen)
- Strawson, P. F., 1980/1994, 'The First Person and Others' in Cassam (ed.) *Self-Knowledge* (Oxford: Oxford University Press)
- Strawson, P. F., 1950, 'On Referring' *Mind* 59: 320-344
- Synofzik, M., Vosgerau, G., and Newen, A., 2008, 'Beyond the comparator model: a multifactorial two-step account of agency' *Conscious and Cognition* 17: 219-239
- Taylor, K. A., 2010, 'On Singularity' in Jeshion (ed.) *New Essays on Singular Thought* (Oxford: Oxford University Press)

- Tomberlin, J.E. (ed.) 1983, *Agent, Language and the Structure of the World* (Indianapolis: Hackett)
- Tye, M., 1995, *Ten Problems of Consciousness* (Cambridge: MIT Press)
- Tye, M., 2009, *Consciousness revisited* (Cambridge, MA: MIT)
- Van Straaten, Z. (ed.) 1980, *Philosophical Subjects: Essays presented to P.F. Strawson* (Oxford: Clarendon Press)
- de Vignemont, F., 2012, 'Bodily immunity to error' in Recanati and Prosser (eds.) *Immunity to error through misidentification: New Essays* (Cambridge: Cambridge University Press)
- de Vignemont, F., forthcoming in Mind 'On the multimodality of bodily perception'
- Wegner, D., 2002, *The Illusion of Conscious Will* (Cambridge, MA: MIT Press)
- Wiggins, D., 1967, *Identity and Spatio-temporal Continuity* (Cambridge: Cambridge University Press)
- Wiggins, D., 1976, 'Butler, and the stream of consciousness: and men as a natural kind', in Rorty (ed.) *The Identities of Persons* (Berkeley: University of California Press)
- Wiggins, D., 1980, *Sameness and Substance* (Cambridge, MA: Harvard University Press)
- Wiggins, D., 2001, *Sameness and Substance Renewed* (Cambridge: Cambridge University Press)
- Wittgenstein, L., 1958, *The Blue and the Brown Books* (Oxford: Blackwell)
- Wong, H.Y., 2015, 'On Proprioception in Action: Multimodality versus Deafferentation', forthcoming
- Wright, C., 1998, 'Self-Knowledge: "The Wittgensteinian Legacy"' in MacDonald et al. (eds.) *Knowing Our Own Minds* (Oxford: Calderon Press)
- Wright, C (ed.) 1984, *Frege: Tradition & Influence* (Oxford: Blackwell)